

AD A 054091

DDC FILE COPY

ENGINEERING

FOR FURTHER TRAN

USCIP Report 800



12

UNIVERSITY OF SOUTHERN CALIFORNIA

SEMIANNUAL TECHNICAL REPORT

Harry C. Andrews
Project Director

Covering Research Activity During the Period
1 October 1977 through 31 March 1978

31 March 1978

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

Sponsored by

Advanced Research Projects Agency
Contract No. F-33615-76-C-1203
ARPA Order No. 3119

ORIGINAL CONTAINS COLOR PLATES; ALL DDC
REPRODUCTIONS WILL BE IN BLACK AND WHITE.



IMAGE PROCESSING INSTITUTE

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

NTIS	White Section	<input checked="" type="checkbox"/>
DDC	Buff Section	<input type="checkbox"/>
UNANNOUNCED		
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
DISL	AVAIL.	SPECIAL
A		

USCIP Report 800

12

SEMIANNUAL TECHNICAL REPORT

Covering Research Activity During the Period
1 October 1977 through 31 March 1978

Harry C. Andrews
Project Director
(213) 741-5514

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

31 March 1978

ORIGINAL CONTAINS COLOR PLATES: ALL DDC
REPRODUCTIONS WILL BE IN BLACK AND WHITE.

DDC
RECEIVED
MAY 15 1978
D

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Wright Patterson Air Force Base under Contract F-33615-76-C-1203, ARPA Order No. 3119. Additional support was provided by AFOSR Contract AFOSR-77-3285 and WPAFB Contract F-33615-77-C-1016.

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Image Processing Institute ✓ University of Southern California University Park Los Angeles, California 90007		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	
3. REPORT TITLE 6 IMAGE UNDERSTANDING RESEARCH.		2b. GROUP	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) 9 Semiannual Technical 1 October 1977 through 31 March 1978			
5. AUTHOR(S) (First name, middle initial, last name) 10 Harry C. Andrews (Project Director) <i>rept.</i>			
6. REPORT DATE 11 31 March 1978		7. TOTAL NO. OF PAGES 14 78	7b. NO. OF REFS 76
8a. CONTRACT OR GRANT NO. 15 F-33615-76-C-1203		8b. ORIGINATOR'S REPORT NUMBER(S) F33615-77-C-1016 USCIPI Report 800	
9. PROJECT NO. ARPA Order No. 3119		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
c. 12 181p.			
10. DISTRIBUTION STATEMENT Approved for release: distribution unlimited			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209	
13. ABSTRACT This technical report summarizes the image understanding, smart sensor, and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 October 1977 through 31 March 1978 under Contract Number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office. The image understanding projects reported herein are maturing with symbolic matching, structure location, edge fitting, stochastic texture analysis and SVD feature selection, all being reported upon in some detail. The image processing projects present both new and concluding projects. New projects include double phase binary computer generated holograms and turntable radar imaging via coherent multi-frequency radar return processing. Older projects resulting in successful theoretics and experimental work include a posteriori restoration and perceptual model color image coding. Our on-going smart sensor project is expanding rapidly with old circuits being driven at near real time TV rates and new circuits being designed for 7 x 7 area processing for both enhancement and texture development. The Institute has installed a real time TV solid state refresh monitor and display at ARPA headquarters. This allows recent pictorial results to be made available over the ARPANET. Any and all contractors can make use of this device with software drivers available from the Institute.			

DD FORM 1473
1 NOV 65

UNCLASSIFIED

Security Classification

391 141

1B

14.

KEY WORDS

LINK A

LINK B

LINK C

ROLE

WT

ROLE

WT

ROLE

WT

Key Words: Digital Image Processing,
Image Restoration, Degrees of Freedom,
Scene Analysis, Image Understanding, Edge
Detection, Image Segmentation, CCD Arrays,
CCD Processors.

UNCLASSIFIED

Security Classification

Abstract

This technical report summarizes the image understanding, smart sensor, and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 October 1977 through 31 March 1978 under Contract Number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The image understanding projects reported herein are maturing with symbolic matching, structure location, edge fitting, stochastic texture analysis and SVD feature selection, all being reported upon in some detail. The image processing projects present both new and concluding projects. New projects include double phase binary computer generated holograms and turntable radar imaging via coherent multi-frequency radar return processing. Older projects resulting in successful theoretic and experimental work include a posteriori restoration and perceptual model color image coding. Our on-going smart sensor project is expanding rapidly with old circuits being driven at near real time TV rates and new circuits being designed for 7×7 area processing for both enhancement and texture development. The Institute has installed a real time TV solid state refresh monitor and display at ARPA headquarters. This allows recent pictorial results to be made available over the ARPANET. Any and all contractors can make use of this device with software drivers available from the Institute.

PROJECT PARTICIPANTS

Project Director

Affiliation

Harry C. Andrews

Computer Science & Electrical
Engineering

Research Staff

Affiliation

Nasser E. Nahi

Electrical Engineering

Ramakant Nevatia

Computer Science

William K. Pratt

Electrical Engineering

Keith E. Price

Image Processing Institute

Alexander A. Sawchuk

Electrical Engineering

Timothy C. Strand

Image Processing Institute

Support Staff

Gerard Ashton

David Nagai

Howard Byus

Clay Olmstead

Marilyn Chan

Beverly Sanders

David Drake

Ray Schmidt

Gary Edwards

Pat Stoliker

John Horner

Jim Tertocha

Chung-Kai Hsueh

Roger Tertocha

Eileen Jurak

Rose Ward

Ed Kasanjian

Steven White

Toyone Mayeda

Amy Yiu

Students

Ikram E. Abdou

Ahmad Armand

Behnam Ashjari

K. Ramesh Babu

Pradeep Bhadsavle

Bir Bhanu

Chung-Ching Chen

Peter Chuan

David Garber

Charles F. Hall

Bijan Lashgari

Kenneth I. Laws

Chun Moo Lo

John Morton

TABLE OF CONTENTS

	Page
1. Research Overview.....	1
2. Image Understanding Projects.....	2
2.1 Matching Segments of Images	
- Keith E. Price.....	2
2.2 Symbolic Matching and Analysis with Substantial	
Changes in Orientation	
- Keith E. Price.....	22
2.3 Locating Structures in Aerial Images	
- Ramakant Nevatia and Keith E. Price.....	41
2.4 A New Edge Fitting Algorithm	
- Ikram E. Abdou.....	59
2.5 Stochastic Texture Analysis	
- William K. Pratt.....	65
2.6 Singular Value Decomposition Feature Extraction	
- Behnam Ashjari and William K. Pratt.....	72
3. Image Processing Projects.....	90
3.1 Double Phase Holograms, A New Way of Generating	
Binary Holograms	
- Chung-Kai Hsueh and Alexander A. Sawchuk.....	90
3.2 A Technique of A Posteriori Restoration --	
Results of a Computer Simulation	
- John Morton.....	105
3.3 Turntable Radar Imaging	
- Chung-Ching Chen and Harry C. Andrews.....	114
3.4 Perceptual Model Coding	
- Charles F. Hall and Harry C. Andrews.....	135
4. Smart Sensor Projects.....	142
4.1 Charge Coupled Device Technology for Smart	

Sensors	
- Graham R. Nudd, Paul A. Nygaard, and Gary D. Thurmond.....	142
4.2 Statement of Work for Follow On CCD Circuitry	
- Harry C. Andrews.....	158
5. Hardware Activities.....	164
5.1 Hardcopy Acquisition	
- Harry C. Andrews.....	164
5.2 The RTTV at ARPA	
- Harry C. Andrews.....	166
6. Recent Ph.D. Dissertations.....	168
6.1 Digital Color Image Compression in a Perceptual Space	
- Charles F. Hall.....	168
7. Recent Institute Personnel Publications.....	169

1. Research Overview

The past six months have been quite productive on a variety of research and development fronts. The image understanding projects are maturing with symbolic matching, structure location, edge fitting, stochastic texture analysis and SVD feature selection, all being reported upon in some detail. The image processing projects present both new and concluding projects. New projects include double phase binary computer generated holograms and turntable radar imaging via coherent multi-frequency radar return processing. Older projects resulting in successful theoretic and experimental work include a posteriori restoration and perceptual model color image coding. Our on-going smart sensor project is expanding rapidly with old circuits being driven at near real time TV rates and new circuits being designed for 7 x 7 area processing for both enhancement and texture development. The Institute has recently acquired a high precision hardcopy color device for improved output capability and has installed a real time TV solid state refresh monitor and display at ARPA headquarters. This allows recent pictorial results to be made available over the ARPANET. Any and all contractors can make use of this device with software drivers available from the Institute. Finally this past six months has witnessed the graduation of one Ph.D. student and numerous Institute personnel publications listed at the end of this semi-annual report.

2. Image Understanding Projects

The projects reported herein describe progress on a variety of fronts in our image understanding efforts. Segment matching is discussed in some detail followed by symbolic matching of images with substantial changes in orientation. Results in both these areas are encouraging as the reader will see in reading these contributions. A technique is next presented for locating structures in aerial images followed by the presentation of a new edge fitting algorithm. A quantitative analysis tool is developed to more accurately define what and where an edge is. These contributions are followed by a discussion on stochastic texture analysis in which some previously accepted perceptual moment assumptions are challenged. Finally the section closes with a discussion on singular value decomposition image feature extraction as a useful tool for obtaining features in an image understanding scenario.

2.1 Matching Segments of Images

Keith E. Price

Introduction

Image registration has few direct applications by itself, but it is a necessary step for computer change analysis, tracking with multiple images, or motion analysis. We define registration as: finding corresponding points or regions in two or more images of a single scene, where the images may be taken from a different position or at a different time. For some registration and change analysis systems, one image may be transformed so that each point in

one image precisely corresponds with the same point in the other image, but this is not always necessary.

There are two basic types of registration methods: signal based and symbolic. Signal based registration systems find pairs of corresponding points by matching the image values in a small neighborhood around the two points using, for example, a correlation function [1-6]. A symbolic registration system locates pairs of corresponding segments by comparing values of features over entire segments [7-10]. The symbolic representation of an image contains segments representing objects or part of objects and a feature based description of the segments.

Signal based registration systems must make a major assumption about the allowable pairs of images: the two images are already in close alignment and major portions of the scene remain unchanged. The first is necessary to limit the area of the image where corresponding points may be located to reduce the search time, and the second assures that enough corresponding points for a good match will look the same in both images. The first restriction can be relaxed if the approximate camera positions are known so that one image can be transformed before any other processing is attempted. These assumptions are valid for controlled stereo pairs and thus most stereo analysis can be performed very well at the signal level. Image based matching is not always possible for arbitrarily selected points if the point falls in the middle of a homogeneous region, thus some processing to locate edges or other features of regions would be necessary in a general system [6].

Analysis after registration and some change analysis is usually best performed using symbolic techniques, e.g.

shadow detection cannot be performed on a single pixel, but requires the analysis of nearby pixels to determine if there might be a shadow and the location of an object that could cast a shadow in that position. Thus it would seem practical to attempt to perform the initial registration symbolically so that the change analysis and later processing can more easily be performed symbolically.

Related work has been reported in [8] and [9] where a limited feature based description is used to compare images of a moving scene and maintain a description of the scene. The description is limited to features invariant to rotations and translations, and is also limited by the fact that the scenes are composed of white objects on a black background. This system uses a model of the objects in motion in the scene to predict the current view so that it can match scenes with occlusions. The assumptions made and problems studied in [9] differ from those in the work reported here so the results are not directly comparable, but both represent similar techniques being applied to different aspects of the image matching problem.

We will first describe the problem which we are attempting to solve and present the scenes which will be used to illustrate important aspects of the solution. The third section will describe the procedures which we have implemented to solve these problems. Section IV will present some of the results from these programs.

Problem Description

We are exploring the general problem of finding changes between different views of the same scene. The views may be taken at different times or from different positions, and the changes can occur in any of features used to describe

objects in the scene - size, shape, color, location, etc. As a necessary step in locating these differences we have worked on a system to locate corresponding segments in two views of one scene. The differences, in many cases, are then simply the differences in the feature values of the two corresponding segments. Several assumptions have been made about the type of input and the extent of changes which occur. First, the images are arbitrary views of natural scenes, including ground level views and aerial photographs, and have already been segmented by a computer program such as in [7, 11-13]. This implies that the segmentations are not perfect and may vary from one view to the next. Second, objects may appear in only one view, and many objects may be very similar to each other. Also, there may be both global and local changes in the scene and changes can occur in any feature, but most corresponding region pairs will be more similar than noncorresponding region pairs. Finally, included in the descriptions of the tasks given to this program is an indication of which features will usually remain constant and which features will usually change. The uses of these assumptions should become apparent in the next section.

Scene Descriptions

Figures 1 through 4 show the initial images and partial segmentations for an urban scene. There are size and position changes for the objects between the two images but no rotation differences. There are many local differences between the images and these are discussed further in [7]. In this study we are interested only in locating corresponding positions in the two images and determining the scale difference. This task can be accomplished best by using several well defined, easily segmented regions, therefore only the bright regions have been segmented in

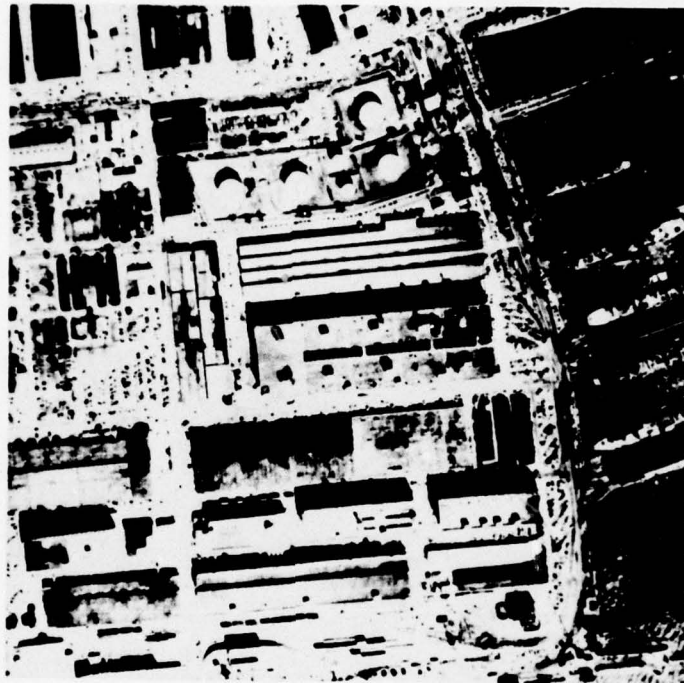


Figure 1. Urban scene, image 1.



Figure 2. Urban scene, image 2.



Figure 3. Urban scene, segmentation of Image 1
The segmented regions are the gray areas
in the figure.

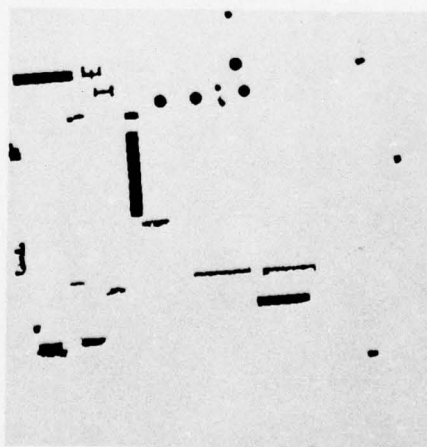


Figure 4. Urban scene, segmentation of Image 2

this pair of images.

Figures 5 through 8 give the input images and segmentations for a rural scene. These images are slightly rotated with respect to each other but were taken at about the same time. The two segmentations include only the large untextured (homogeneous) regions and the much smaller bright regions. This scene illustrates the operation of the matching procedure when there is a small orientation difference and when there are several similar regions extracted in the segmentation operation. These two are also discussed further in [7] along with the description of the operation of the segmentation procedure as applied to both of these scenes.

Matching

In a system which is designed to match a pair of real images, the problem is usually not whether the segments match exactly, using a given feature, but how close the segments match. When applied to the problem of finding the corresponding region in a set of regions, the question becomes: how well do these two regions match compared to how well the region in the first image matches other regions in the second image? Thus we want a match function which will indicate how well two segments match by combining how well segments match for each feature.

A first approximation of how well two regions match for one feature is given by the difference between the feature values. This might be reasonable if all features had similar values (e.g. only spectral features), but when combining arbitrary features the absolute difference is not meaningful. Therefore, the feature value differences must be weighted so that the contribution to the matching



Figure 5. Rural scene, image 1.

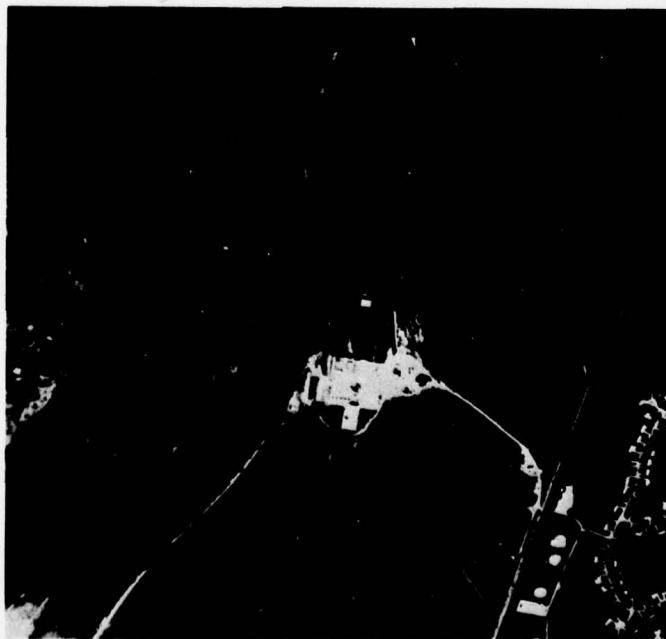


Figure 6. Rural scene, image 2.

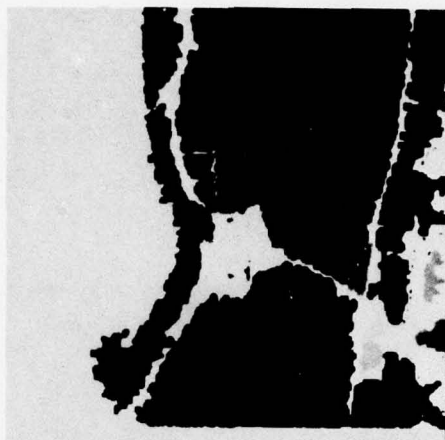


Figure 7. Rural scene, segmentation of Image 1
Each of the large regions is displayed with a different gray level and the small bright regions are given one gray value.



Figure 8. Rural scene, segmentation of Image 2

function of each feature is approximately the same for equally poor matches. This feature normalization value must account for several variations: the sizes of the feature values are very different - some range from 0 to 1, some from 10 to one million (size); and some features vary more with small differences in the image than others or are more prone to errors in computation. Table 1 gives a list of the features and the weighting for the actual matching procedure. We chose a simple method to combine the weighted differences, just the sum of the weighted absolute differences. Thus the region to region match rating is:

$$R_{ij} = - \sum_{k=1}^n |V_{ik} - V_{jk}| W_k \quad (1)$$

where R_{ij} is the value of the match between regions i and j , n is the number of valid features, V_{ik} (V_{jk}) is the value of the k th feature of regions i (j), and W_k is the normalization weight for the k th feature. Region i and j are from different images. The result is negative so that better matches have a greater value, i.e. closer to zero.

The neighbor and relative position features have no affect until at least one pair of corresponding regions is located. The neighbor feature value for region i and j are computed as follows: Let N_i be the set of all regions which are neighbors of region i and already have a corresponding region in the second image. Let N_j be the set of the corresponding regions for those regions in N_i . Then let N'_j be the members of N_j which are neighbors of region j . The neighbor feature value for region i is the length of N_i and for region j is the length of N'_j . The relative position feature (above, below, to right, and to left) value is defined in an equivalent manner.

<u>Feature</u>	<u>Feature Weighting</u>
Size	Minimum of: $5/\text{Area of region,}$ $100/\text{Area of image}$
Location, each coordinate	$8/\text{Image dimension}$
$\text{Perimeter}^2/\text{Area}$	$2A/P^2$
Orientation	2
Length to Width Ratio	2
Area/Area of Minimum Bounding Rectangle	1.6667
Area/Area of Bounding Ellipse	2.5
Color or Intensity	$1/2\sigma$ Standard deviation of the color parameter over the given region.
Neighbors	$2/\text{Number of neighbors with a}$ corresponding region
Relative Positions	$2/\text{Number of relative regions with a}$ corresponding region

Table 1. FEATURE VALUE NORMALIZATION WEIGHTS. These are the weights used in Eqs. (1) and (2) for all the available features. These values are used to make the contribution of each feature to the match rating approximately equal. When feature values are indicated in the definition of the weighting, the values are from image 1 or the region in image 1. The features with constant weight have values in a limited range, such as 0 to 1.

We stated in the problem description section that the task description may include the fact that some features are changing and are thus unreliable for use in the matching process. Alternatively, we may be given the fact that some feature must not change between the two views and these features are very reliable. Therefore, for a given task, some features should contribute more to the region to region match rating than other features and some features may not even be considered. We added a strength parameter to the rating function to account for the different types of features. This gives a region to region match rating of:

$$R_{ij} = - \sum_{k=1}^n |V_{ik} - V_{jk}| W_k S_k \quad (2)$$

where S_k is the strength of the k th feature. There are only three possible strength values which are used by the program. These values were selected so that it requires several poor matches in lower strength features to have the same impact as an equally poor match in one higher strength feature. The features are assigned to strengths based on information given in the task description, with features assigned the medium strength when the outside information does not specify otherwise. If a good match in a feature is indicated as absolutely necessary then the feature is given the higher strength and if it is given that the feature may change, but not too dramatically, then the lower strength is used. Features which are not available and those which are known to be totally unreliable for this task are not considered in the rating - effectively a fourth strength of zero. Generally, the assignment of strengths is the same for all matches between a pair of images.

Certain global changes may mean that some features cannot be used in the matching operation, for example scale

changes mean that size, and absolute position will not match. But, adjustments in the feature values can be made if these global changes are known. We indicated earlier that knowledge of which properties are changing may be given, but not the magnitude of these changes, which would be necessary to make any adjustments. But detailed information about the global changes is available, if several corresponding region pairs can be located without using these changing features. After the adjustments are computed the feature can be used like any other feature, except that the feature values will be corrected before they are used in the matching function. In this case, the strength assignments will change after the first few matches because the adjusted feature values will match better than the original feature values.

This region matching process is the basic operation for the procedure which locates corresponding regions in two images. The symbolic registration procedure compares the given region with each potential matching region using given sets of features with different strengths (e.g. those which must remain constant and those which may change). The region which best matches the given region is returned as the corresponding region. It should be noted that this is not an attempt to find a complete one to one mapping from regions in one image to regions in another image, since this is not possible when objects may appear or disappear and different regions may be segmented in the two images. The regions to be registered and the strengths assigned to each feature are not selected by the symbolic registration procedure, but are given by other elements of the total system [7]. The order in which the corresponding regions are located may be important. It is especially important that initial matches be accurate if the neighbor and relative position features are to be of any value. The

usual order is to find matches for the regions in order of size so that the neighbor and relative position features can be used. But when a few initial matches are needed to make adjustments for global changes, the large regions may provide unreliable information especially if they are adjacent to the edge of the image. Because of this, easily identified regions, other than the larger regions, may also be matched first, but these decisions are made a controlling system or the user, and are based on the task description.

Results

Figures 9 and 10 show the registration results for the two scenes, i.e. the corresponding regions in the two images - the corresponding regions have the same intensity values in the two pictures which make up each figure. Since only the bright regions were segmented in the urban scene, these are the only ones which are used in the registration process. This pair of images has a global scale and position change which is computed after the first few corresponding regions are located (the large horizontal object in the lower right and the left most round object in the upper center). These adjustments are used for all subsequent matching operations so that location and size can also be treated as constant features. This adjustment is necessary to obtain correct matches for many of the regions, especially the group of round regions near the top of the scene where the shape of the extracted segments differ because of shadows which make position an important features in the matching operation. The computed global size change is about 1.5.

The large regions in the rural scene images are large homogeneous (nontextured) regions and there are many differences in the segmentations of these regions because

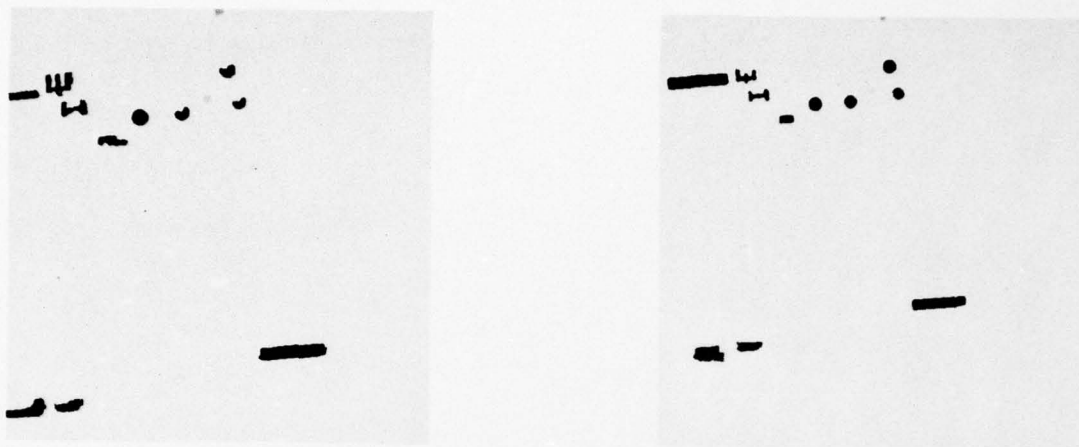


Figure 9. Symbolic Registration for urban scene
The corresponding regions have the same gray value in the two pictures. Image 1 is on the left and image 2 on the right. Some of the small round regions may not be visible since their values are almost white.



Figure 10. Symbolic Registration for rural scene
Corresponding regions have the same gray level in the two pictures. Image 1 is on the left and image 2 on the right. The bright areas near the center are almost white and some may not be apparent in the pictures.

this feature is very sensitive to noise. Because of this, there are some regions which have no correct corresponding region among the segmented regions in the other image. But generally, when they exist, the correct corresponding regions are located. The results for the smaller regions might be improved if an adjustment for the rotation difference is calculated and used, but many of the houses are correctly matched even though they are all similar to each other.

One immediate question that may be raised is how close the best match rating is to ratings for other potential matches. The mean rating for the best match for the rural scene shown here is -287 ($\sigma=116$) and the mean for the second best matching region (i.e. the region which would have been selected if the actual matching region had not been segmented) is -530 ($\sigma=301$). Values similar to these are common for other scenes which we have analyzed. In most cases the mean for the second best matching pairs is about one standard deviation from the mean for the best matching pairs (in some cases it is much further, but rarely much closer). Tables 2 and 3 show more detailed region to region matching results for the urban and rural scene respectively. Regions in the first image correspond to rows in the table, and those in image 2 to columns. The values are the region to region match values for generated by comparing regions in image 1 to regions in image 2. The values along the diagonal are the match values for the computed corresponding regions. The regions in each table are ordered by the size of the region in the first image. There is a wide disparity between the values of accepted matches, from -77 to -590. The differences between the rating for the accepted pair and other pairs can be very small when the regions are almost identical such as the round regions in the urban scene, noted by *, and the houses in the rural scene, marked by +.

-77	-1369	-925	-1251	-791	-1285	-1166	-1156	-1200	-1201	-1107	-1281	-1182	-1125
-1736	-410	-1436	-741	-1359	-1061	-569	-925	-1244	-1139	-1252	-1258	-1156	-1249
-1002	-1150	-300	-1075	-366	-1416	-994	-1182	-1712	-1517	-1670	-1768	-1606	-1574
-1500	-561	-1126	-277	-1189	-984	-467	-808	-1235	-1060	-1203	-1211	-1108	-1184
-874	-1120	-430	-1101	-119	-1251	-867	-989	-1549	-1350	-1505	-1590	-1445	-1414
*-1637	-1222	-1485	-1113	-1405	-138	-968	-537	-564	-272	-473	-571	-433	-406
-1379	-383	-1201	-548	-1013	-743	-118	-610	-1071	-843	-1002	-1092	-921	-968
-1010	-511	-984	-521	-832	-537	-333	-316	-836	-636	-790	-874	-727	-695
*-1307	-1151	-1673	-1085	-1505	-720	-1023	-699	-321	-617	-527	-480	-542	-523
*-1495	-1243	-1598	-1126	-1493	-473	-992	-663	-682	-340	-561	-653	-501	-485
*-1391	-1319	-1754	-1257	-1632	-718	-1145	-803	-580	-614	-377	-620	-528	-535
-1965	-1653	-2130	-1639	-2072	-1067	-1633	-1234	-853	-951	-922	-379	-872	-932
*-1510	-1304	-1671	-1189	-1576	-387	-1126	-703	-317	-289	-286	-367	-156	-256
*-1576	-1501	-1727	-1357	-1653	-396	-1317	-745	-390	-263	-253	-458	-290	-136

Table 2. Region to region match results for the urban scene.

The regions are ordered by the size of the region in image 1. The numbers in a given row are the match values for one region in image 1 with each region in image 2. The diagonal entries are the values for the computed corresponding regions. The "*"s indicate the group of round regions.

21	96	206	158	90	158	90	157	133	133	145	209	227	151	166	180	195	195	180	168	191	191	186	191	195
129	12	194	118	128	118	128	126	117	117	136	188	230	160	151	182	155	173	182	162	185	169	165	161	155
306	166	48	176	135	176	135	185	167	167	156	131	157	194	136	212	211	193	212	193	205	216	196	193	211
318	201	226	41	132	41	132	64	66	66	143	178	209	159	156	124	93	120	124	109	132	111	113	111	93
310	243	163	156	59	156	59	173	155	155	135	179	196	152	124	204	209	202	204	185	201	213	198	200	209
308	213	229	49	89	49	89	56	40	40	107	151	186	130	118	93	98	93	93	74	97	96	89	90	98
305	216	174	110	22	110	22	118	104	104	82	142	163	100	83	151	157	153	151	131	150	161	149	150	157
330	223	274	51	113	51	113	29	31	31	113	172	205	131	132	83	80	82	83	69	88	67	70	72	80
337	204	301	59	107	59	107	55	40	40	100	182	221	114	123	88	66	94	88	84	93	73	83	83	66
332	208	307	65	105	65	105	57	32	32	94	179	218	103	118	76	66	83	76	77	82	62	72	72	66
374	223	219	126	106	126	106	131	111	111	37	64	111	69	46	89	91	88	89	68	92	88	84	86	91
393	251	180	156	122	156	122	157	140	140	51	21	50	84	52	102	105	89	102	83	95	107	93	90	105
382	263	156	160	109	160	109	158	144	144	57	33	27	61	50	98	104	81	98	93	94	107	89	85	104
370	253	312	161	94	161	94	153	127	127	45	149	169	18	67	64	93	101	64	88	76	82	93	100	93
400	255	182	156	122	156	122	161	145	145	47	47	78	88	24	109	109	90	109	87	103	113	93	90	109
+364	285	365	132	138	132	138	109	92	92	77	177	197	66	100	26	58	55	26	44	35	39	49	58	58
+393	261	324	99	155	99	155	75	86	86	102	144	164	104	107	54	32	35	54	52	45	35	32	30	32
+369	269	269	105	143	105	143	96	88	88	87	110	126	95	75	51	46	19	51	39	44	51	25	25	46
+362	279	352	127	140	127	140	109	88	88	75	169	194	62	99	21	54	58	21	43	37	38	49	58	54
+352	267	283	109	132	109	132	98	87	87	65	112	139	87	73	44	55	44	44	24	48	54	40	43	55
+383	293	379	144	147	144	147	119	106	106	84	182	206	74	100	26	54	54	26	44	19	38	46	56	54
+375	274	364	119	151	119	151	88	75	75	86	174	204	77	111	20	45	46	20	36	28	15	39	49	45
+370	263	287	105	141	105	141	93	84	84	83	124	144	90	82	47	38	25	47	36	40	44	17	24	38
+384	269	287	106	154	106	154	101	91	91	90	115	131	100	80	59	39	15	59	43	51	49	18	14	39
+397	262	342	99	157	99	157	90	87	87	97	163	188	93	108	52	23	45	52	54	47	38	31	34	23

Table 3. Region to region match results for the rural scene.

To reduce the size of this table all numbers have been divided by 10 and the minus signs have been removed, otherwise the format is the same as Table 2. The "+"s indicate the regions which correspond to houses in the scene.

In some cases two regions in the first image match with the same region in the second image so some columns are equal.

Conclusions

This work is intended to be part of a general purpose analysis system and not simply the solution to a particular set of problems. The symbolic registration procedure worked well with the machine segmentations and performance should improve if segmentations were improved. This research shows that symbolic techniques can be used for analysis of a larger class of scenes than is possible with signal based system, but may be used in conjunction with signal based systems where detailed analysis of changes in individual pixels is required. Since many change analysis tasks are best performed at a symbolic level, it seems reasonable to perform the matching and registration symbolically rather than generate symbolic descriptions of unreliable difference images, thus symbolic registration will be an important step of general change analysis systems.

References

1. L.H. Quam, "Computer Comparison of Pictures," Ph.D. Thesis, AIM-144, Stanford University, Stanford, California, May 1971.
2. R.L. Lilestrand, "Techniques of Change Detection," IEEE Transactions on Computers, July 1972.
3. G.R. Allen, L.O. Bonrud, J.J. Cosgrove, and R.M. Stone, "The Design and Use of Special Purpose Processors for the Machine Processing of Remotely Sensed Data," IEEE Symposium on Machine Processsing of Remotely Sensed Data, Purdue University, October 1973.

4. M.D. Levine, D.A. O'Handley, and G.M. Yagi, "Computer Determination of Depth Maps," Computer Graphics and Image Processing, No. 2, 1973, pp. 131-150.
5. M.J. Hannan, "Computer Matching of Areas in Stereo Images," Stanford Artificial Intelligence Laboratory Memo 239, Stanford University, Stanford, California, July 1974.
6. H.P. Moravec, "Towards Automatic Visual Obstacle Avoidance," in Proc. IJCAI-77, Cambridge, Ma., 1977, p. 584.
7. K. Price, "Change Analysis and Detection in Multi-Spectral Images," Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, December 1976.
8. J.K. Aggarwal and R.O. Duda, "Computer Analysis of Moving Polygonal Images," IEEE-TC 24, 1975, pp. 966-76.
9. W.K. Chow and J.K. Aggarwal, "Computer Analysis of Planer Curvilinear Moving Images," IEEE-TC 26, 1977, pp. 179-185.
10. K. Price, R. Reddy, "Change Detection and Analysis in Multispectral Images," in Proc. IJCAI-77, Cambridge, Ma., p 177, pp. 619-625.
11. R. Ohlander, K. Price, R. Reddy, "Picture Segmentation Using a Recursive Region Splitting Method," Computer Graphics and Image Processing, to appear.
12. J.M. Tenenbaum, T.D. Garvey, S. Weyl, H.C. Wolf, "An Interactive Facility for Scene Analysis Research," Artificial Intelligence Center Technical Note 87, Stanford

Research Institute, Menlo Park, California, January 1974.

13. Y. Yakimovsky, "Scene Analysis Using a Semantic Base for Region Growing," Ph.D. Thesis, AIM-209, Stanford University, Stanford, California, July 1973.

2.2 Symbolic Matching and Analysis with Substantial Changes in Orientation

Keith E. Price

Introduction

Several methods have been developed which can be used to find correspondences in pairs of images of a changing scene [1-6]. But, for various reasons, these systems did not operate on images which had major changes in orientation - unless an approximate value for the difference was known a priori. Image based methods [1-3] were not used to attempt a solution to this problem because of the complexity of searching for corresponding points. Other work [5,6] assumes that the views are close - in time and position - so that major changes in the point of view are not possible. But in a more general image matching and analysis system [4], this type of problem must be considered.

We have undertaken a series of experiments to examine whether the matching techniques described in [4] can be easily applied to pairs of images which have substantial global differences. In this paper, we present the results of applying this basic technique on pairs of images with orientation differences of 45° , 90° and 180° . These orientation changes are in addition to other, less drastic,

changes which may occur between the two views. The point is to show that symbolic techniques can be applied when there are substantial global changes in addition to the normal local changes which occur between two views of one scene.

We will first discuss the images which will be used for this experiment and describe the results which are desired. Then we will present an outline of the symbolic matching procedure [4] and some initial results using this method. The results suggest some modifications which are then described, followed by more extensive results using the modified procedure.

Tasks for Matching

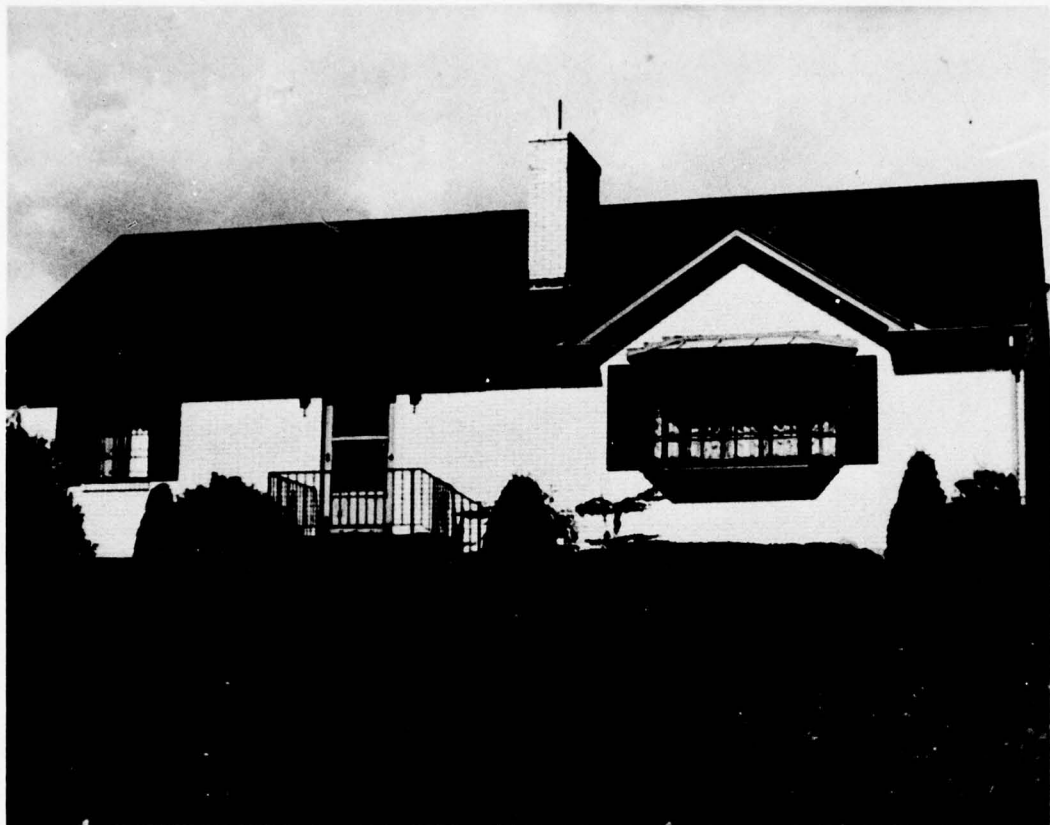
The pairs of images which we will use here have been used earlier [7], but in the previous analyses there were no substantial global changes. We start with a pair of images of a scene taken from slightly different positions, and generate the orientation changes by rotating the digital representations of the second image of the pair. The original first image and the rotated second images are then processed to generate symbolic descriptions (a segmentation into distinct parts plus a description of the segments) which are used by the matching procedures. The details of the segmentation and description are given elsewhere and are not important for this paper [4,7,8].

The basic task to be executed for the images presented here is to find the corresponding regions in the two images when there is a large, but unknown, orientation difference. A by-product of this matching should be some indication of the actual orientation difference. The images in this experiment were rotated 45° , 90° , and 180° , but the features which are orientation dependent are not used even if they

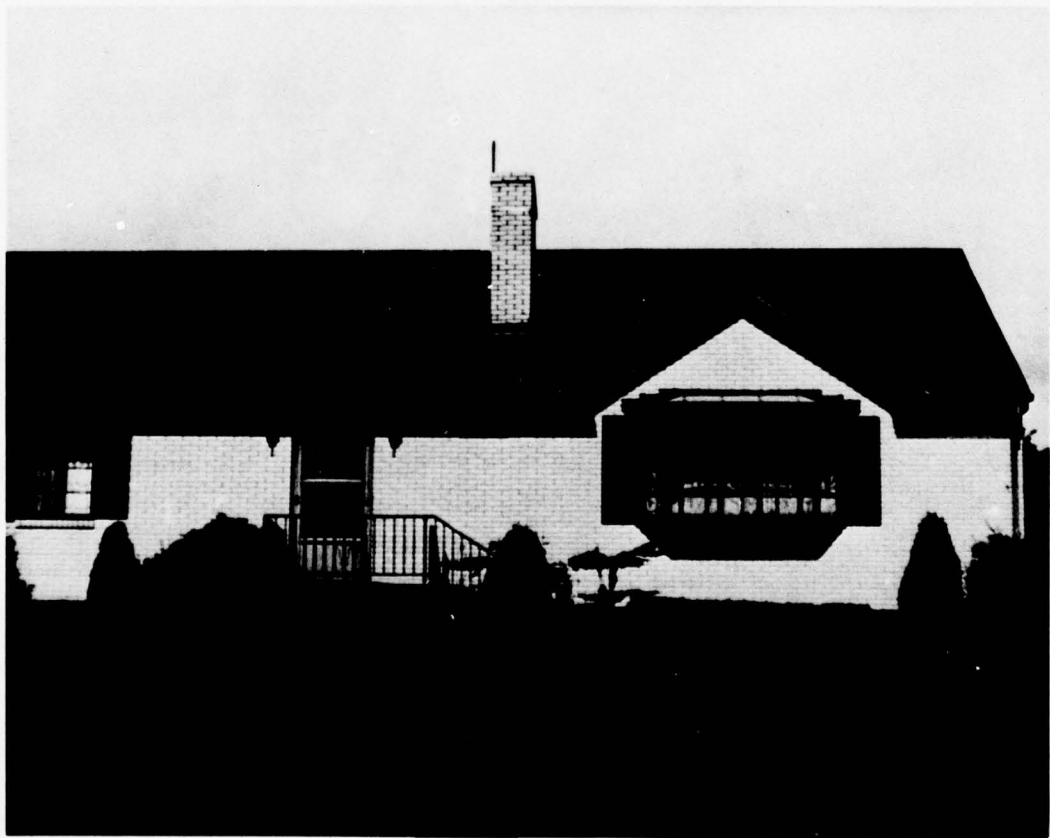
are independent to rotations of 90^0 (e.g. ratio of area and area of minimum bounding rectangle).

The first pair of images, a house, is shown in Figure 1. These are color images, shown here in black and white, so there are several spectral features available for use in the segmentation, description, and matching operations. Figure 2 gives the segmentations of the original first view and the three rotated second views. There are some differences between the regions segmented in the first view and each of the second views, and a few differences among the three second views. The images are not square so the display program puts a white band on the right or bottom, depending on which dimension is smaller. Additionally unsegmented areas are displayed as white areas.

The second pair of images is shown in Figure 3. These are side looking radar views and there is only one spectral input, so that a good segmentation is more difficult and the description is less detailed than the preceding house scene. The important changes in this scene are in objects which are too small to be segmented by our general segmentation techniques but might be easily located by special methods. Figure 4 gives the four segmentations which will be used. There are a few large regions segmented in the first view which are not segmented in the second views. These differently textured regions vary in size and appearance between the two views and would not be used in the matching, therefore they were not segmented in the second views. The regions which are segmented are the untextured areas, all uniformly dark, some of which appear the same in both views. The locations of these few regions can be used to determine the global changes and could aid another system in locating the detailed changes.



(a)

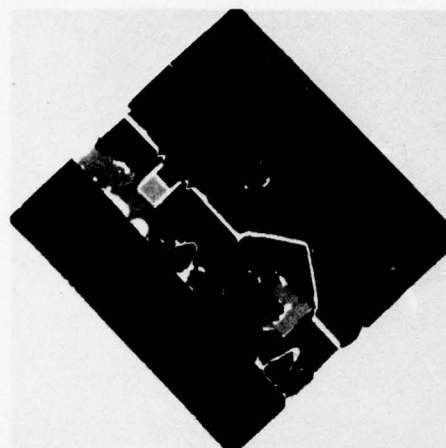


(b)

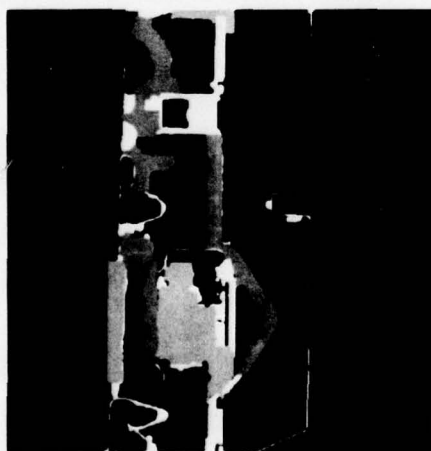
Fig. 1. House images.



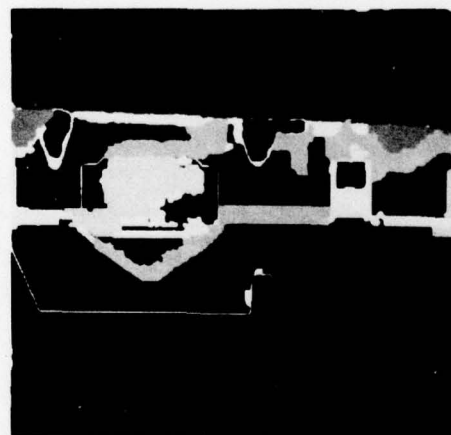
(a) First view.



(b) Second view rotated 45° .



(c) Second view rotated 90° .



(d) Second view rotated 180° .

Fig. 2. Segmentations of house images.



(a)



(b)

Fig. 3. Radar images.



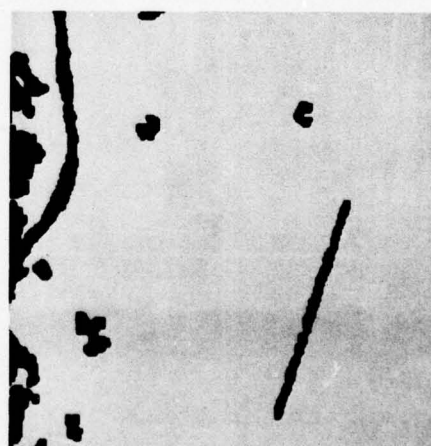
(a) First view.



(b) Second view rotated 45° .



(c) Second view rotated 90° .



(d) Second view rotated 180° .

Fig. 4. Segmentations of radar images.

Matching Procedure Outline

A more detailed description of the matching procedure has been presented elsewhere [4] and will only be outlined here. This matching procedure uses a feature based, symbolic, description of the images. The basic unit in this description is an individual segment generated by an automatic segmentation procedure [8]. These segments are usually regions in the images, but linear features can be described, too. Features which characterize properties such as color, texture, size, shape, position, and adjacencies are used.

The matching procedure is also given an indication of which features are available for matching the current pair of images, and what strength to give to the mismatch using each of these feature. For example, some features are not always computed, red and green in a black and white picture, and some are given as more likely to change than others and are thus given less weight in the matching operation. The match procedure computes a rating for the match between two regions which is essentially a sum of the weighted absolute differences of feature values. The weights used in the sum are composed of a normalization factor to make the contribution from each feature approximately equal and a strength factor to account for the different strengths assigned each feature. There is only a small set of possible strength values, currently 3 different values.

Known global changes between the two images can be used to adjust some feature values, such as size, position and orientation. But, these changes are not given a priori and must be computed from a few initial pairs of matching regions. Thus, the very clearly defined regions should be

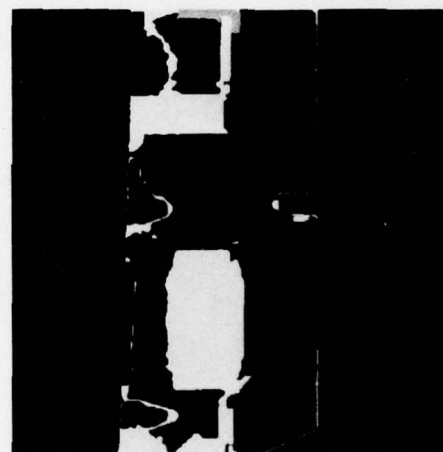


Fig. 5. Initial matching regions for house first view (left) and house second view rotated 90° (right).

matched first so that they may be used for calculating some global changes. In this context, clearly defined regions are regions with extreme values for some feature, e.g. largest, brightest, longest, etc.

We will now present results of applying this procedure to the rotated images and discuss what changes are necessary to achieve accurate results.

Initial Results

Figures 5 and 6 show the results obtained with the above matching procedure for the house scene (90° and 180°). In these, and all other figures, the corresponding regions are displayed at the same intensity in the pair of output pictures. Similar results are obtained for 45° and for matching the second image to the first, but the point here is to point out some of the problems. The orientation feature adjustment was computed, using the sky or roof, and was also used to get these results, but there are still many errors. Most of the unmatched regions would match to an incorrect corresponding region if a match is attempted.

The major problem is that the location of the region is needed to correctly locate matches for many of the smaller regions. This is especially the case when there are size and shape changes due to segmentation differences, such as in the bush, window and door regions. If exact camera transformations are known then the locations in one image can be mapped exactly into locations in the other, but this transformation is not known. We are using many features other than position so an appropriate mapping is sufficient to allow the use of the absolute position features. Given 3 pairs of corresponding regions we can compute a transformation which will map coordinates in one image to

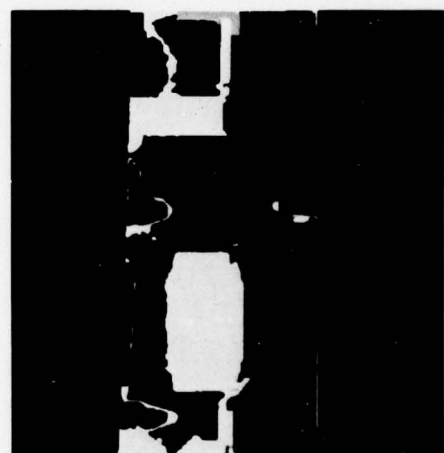


Fig. 5. Initial matching regions for house first view (left) and house second view rotated 90° (right).

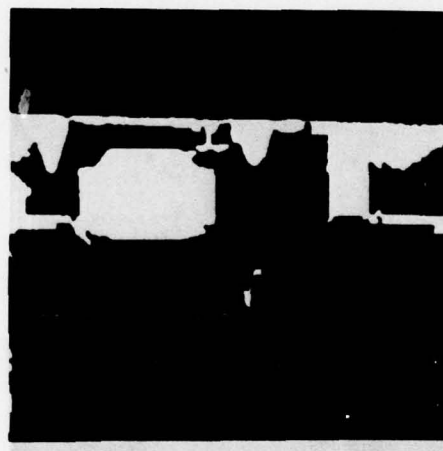
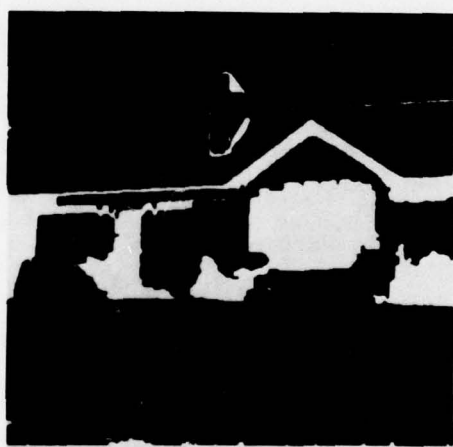


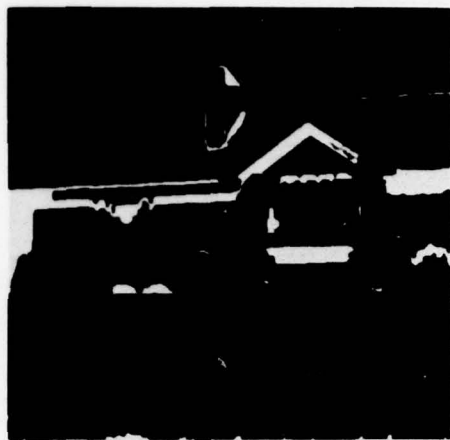
Fig. 6. Initial matching regions for house first view (left) and house second view rotated 180° (right).

coordinates in the other by solving 2 sets of 3 equations and 3 unknowns. This transformation is not optimal for all regions in the image and only accounts for rigid, global changes - e.g. rotations and translations. But this transformation does make the position features useable when there are large global orientation changes.

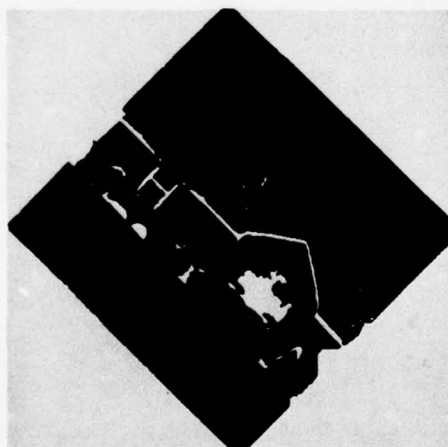
Final Results

Figures 7-9 show the results for the house matching using the computed location transformations - a different transformation is computed for each image pair. Results are given for matching regions in image 1 with image 2 and regions in image 2 with image 1. In this set of images the sky, lawn, and one of the wall sections are the initial 3 regions used for the transformation computation. The transformations do not rotate the coordinates precisely 45° , 90° , or 180° because of differences in segmentations (the sky and lawn are adjacent to the edge and this cause some changes in size and shape) and a small orientation difference which existed before the large rotations were added. But, the adjustment is accurate enough to use the location feature in the match operation. The results at 45° are less accurate than for the other two, but most of the extra mistakes are accounted for by the greater differences in segmentations (see Fig. 2). When two regions in one image correspond to one region in the second image, such as the 2 "bushes" on the right side of the house in the second image appearing as one region in the first image, only one correspondence appears in the output.

Figures 10 and 11 present the results for the radar images. There are very few (i.e. 6) corresponding regions in these two images and two of the pairs have very large size changes. The results for 180° are identical to those

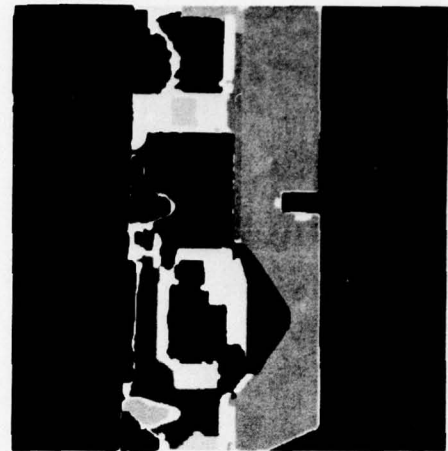


(a) House 1 with house rotated 45° .

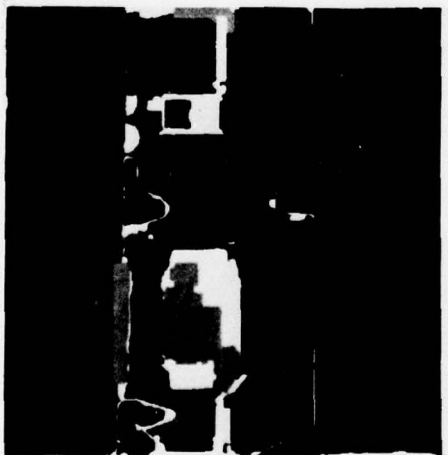


(b) House 2 rotated 45° with house 1.

Fig. 7. Matching.



(a) House 1 with house 2 rotated 90° .

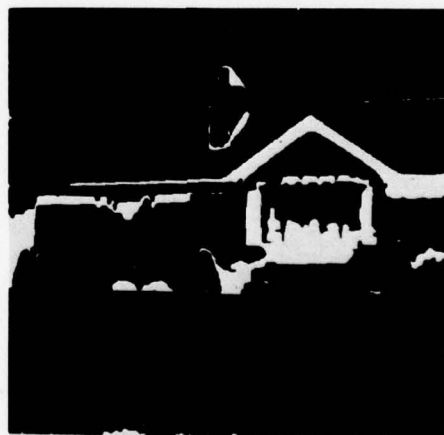
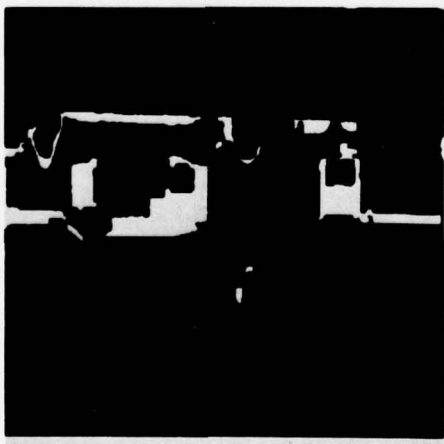


(b) House 2 rotated 90° with house 1.

Fig. 8. Matching

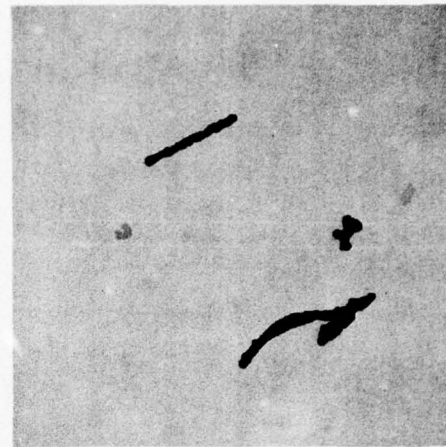


(a) House 1 with house 2 rotated 180° .

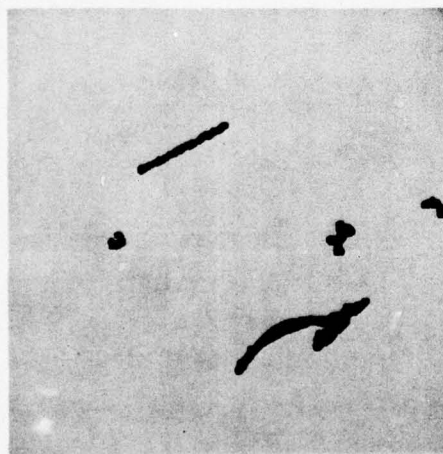


(b) House 2 rotated 180° with house 1.

Fig. 9. Matching

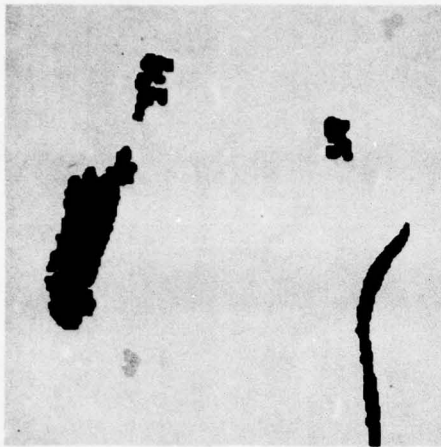


(a) Radar 1 with radar 2 rotated 45° .

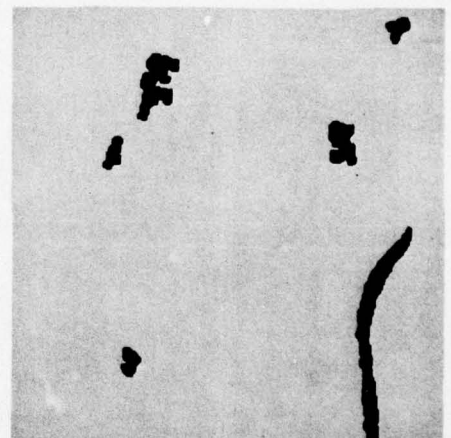
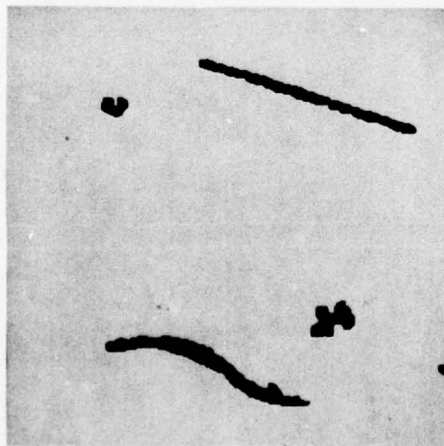


(b) Radar 2 rotated 45° with radar 1.

Fig. 10. Matching



(a) Radar 1 with radar 2 rotated 90° .



(b) Radar 2 rotated 90° with radar 1.

Fig. 11. Matching

for 90° and are not presented. Two of the corresponding pairs for the image 1 to image 2 matching may be difficult to see since they are nearly white in the results picture - one is a correct match (the reversed "C" shape in the lower left), and the other is incorrect (a blob near the top right). The reverse "C" region, the river (lower right) and the blob above the river are the three regions used to compute the transformation in this set of images. This scene shows that when these symbolic techniques are applied to scenes with a reduced feature set - no colors and no neighboring regions, accurate results are possible.

Summary

The complete symbolic registration system presented here has the following basic steps:

1. Segment both images of the scene.
2. Generate a feature based description of the segmented images.
3. Find corresponding regions for the most obvious regions.
4. Set orientation and size correction factors, if necessary.
5. Find several corresponding region pairs.
6. Compute an approximate coordinate transformation, if necessary.
7. Using transformed positions, find all corresponding region pairs.

The matching results depend somewhat on the quality of the segmentation, but the results of these experiments show that this symbolic technique can be used to find corresponding areas in pairs of images even when there are major global changes. We would expect similar, or better, results for pairs of images with global scale, position, and

color changes. We expect less reliable results for scenes with major global changes in all four (orientation, scale, color, and position) because so few features are invariant to all these changes (e.g. relative size, shape measures, and neighbors). But if a controlling system could provide proper guidance, corresponding regions might be located which would account for each of the global changes, separately. For example, scale changes could be based on matching the largest regions, orientation changes might be based on regions with distinctive shape, and so on. But, primary regions with unusual, or extreme, feature values could be used when there are many global changes. In conclusion, symbolic matching methods can work with major global differences, these differences can be detected, and they can be used to great advantage in later analysis.

References

1. G. R. Allen, L. O. Bonrud, J. J. Cosgrove, and R. M. Stone, "The Design and Use of Special Purpose Processors for the Machine Processing of Remotely Sensed Data," IEEE Symposium on Machine Processing of Remotely Sensed Data, Purdue University, October 1973.
2. L. H. Quam, "Computer Comparison of Pictures," Ph. D. Thesis, AIM-144, Stanford University, Stanford, California, May 1971.
3. H. P. Moravec, "Towards Automatic Visual Obstacle Avoidance," in Proc. IJCAI-77, Cambridge, Ma., 1977, p. 584.
4. K. Price, R. Reddy, "Matching Segments of Images," submitted for publication IEEE-TC.

5. W. K. Chow and J. K. Aggarwal, "Computer Analysis of Planer Curvilinear Moving Images," IEEE-TC 26, 1977, pp. 179-185.

6. H. H. Nagel, "Formation of an Object Concept by Analysis of Systematic Time Variations in the Optically Perceptible Environment," Computer Graphics and Image Processing, to appear.

7. K. Price, "Change Analysis and Detection in Multi-Spectral Images," Ph. D. Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, December 1976.

8. R. Ohlander, K. Price, R. Reddy, "Picture Segmentation Using a Recursive Region Splitting Method," Computer Graphics and Image Processing, to appear.

2.3 Locating Structures in Aerial Images

Ramakant Nevatia and Keith E. Price

(Some of this work was reported previously in the last progress report and presented at the October, 1977 ARPA workshop. This material is repeated here for completeness and new results are reported.)

This section describes our initial attempts to utilize our segmentation techniques to perform higher level tasks of locating user specified structures in aerial images. We use edge based [1] as well as region based segmentation techniques [2-3], as experience indicate that the two methods have complementary strengths [4]. The edge techniques are suitable for extracting elongated features

such as roads and the region methods for homogeneous areas such as lakes.

Analysis of aerial images is, in general, a complex task. The reasons for such complexities are many and varied. A prime cause is the presence of texture which causes difficulties for the low level processes such as edge detection and segmentation. Another source of difficulty is that the desired objects and structures may be small compared to the size of a complete image. A detailed analysis of a complete high resolution aerial image is generally prohibitive because of the computational costs.

For many applications, however, a complete and general analysis is unnecessary. Specific structures of interest may have special properties, known a priori, that allow for their easy extraction. The problem of searching for small structures is helped by locating them by their spatial relationships to larger, more easily located structures. We assume that the user specifies the properties useful for locating the desired structures and also the properties of and relations with other objects. An interactive system to facilitate the entry of such data is also being developed. Such a priori information is likely to be available in many applications where different images of the same area need to be analyzed.

Model and Description

The a priori information is represented as a graph structure with the objects represented as nodes and relationships between objects as arcs. Generally, the properties of objects and relationships between objects should be unrestricted. Currently an object is described as a collection of related line segments and regions. Both

lines and regions are described by position (both absolute and relative), size, and orientation features, and regions also have shape (parameter²/area, length/width, area/area of minimum bounding rectangle), color (intensity, red, green, etc.), and texture (e.g. variance) properties. The description and matching operations use both regions and lines with no direct indication of the type of segment needed, because the lines are essentially described as narrow elongated regions.

Several relationships between regions are currently used such as neighbors, nearby regions, and the symbolic relationships of above, below, left and right. Other relationships such as symmetry and similarity are useful but have not been implemented. Simpler relations such as larger, brighter, etc. are not directly implemented, but could be easily computed from given feature values.

The image is described in the same manner as the user specified objects, except that the image description is performed automatically, without direct user input. We will use the term model to mean the descriptions of the user defined object or group of objects which we are attempting to locate in an image.

For example, in Figure 1 we have an aerial view of a mostly rural scene and we wish to locate the docks along the lower side of the river. It might be possible to search the entire image looking for the desired objects, but it seems to be easier to locate the docks by using larger regions and linear features such as lakes and roads. Figure 2 shows a hand drawn pictorial representation of this portion of the image with only the important objects represented. Figure 3 gives the internal graph structure representation of objects which can be used to locate the docks. The a priori



Figure 1. Aerial Image 1.

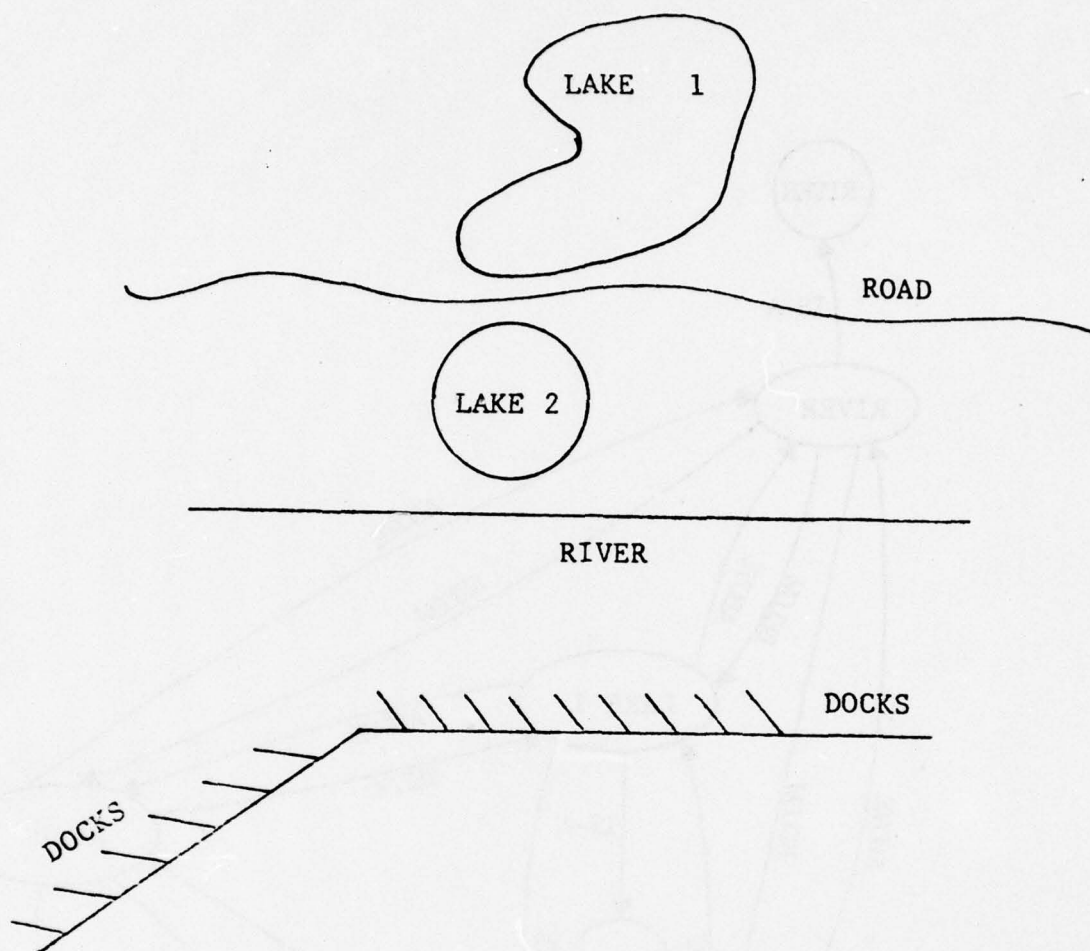


Figure 2. Pictorial Representation of DOCKS Area.

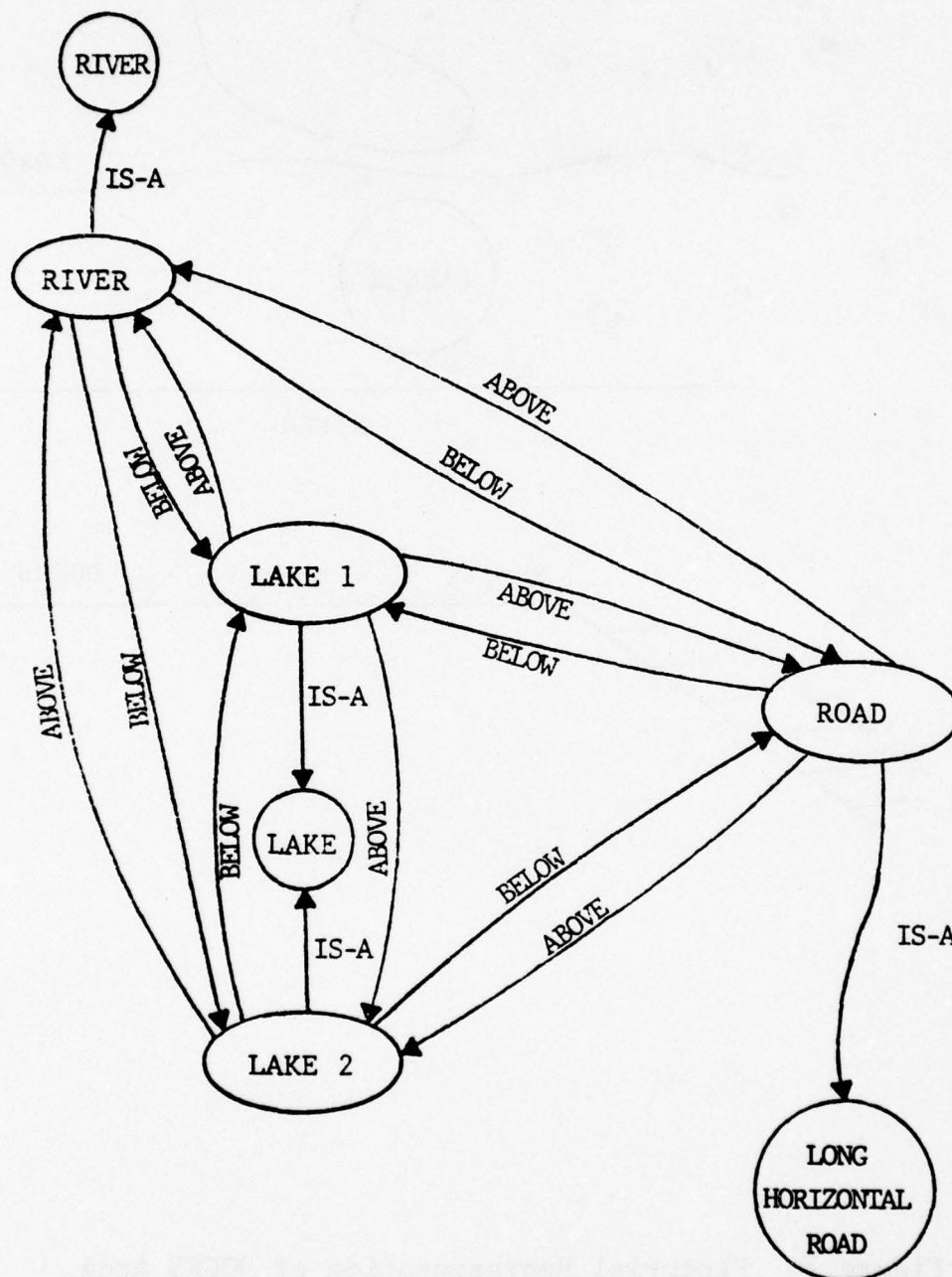


Figure 3. Internal, Graph, Representation of Model.

knowledge must also provide the system with a way of locating the area containing the docks given the positions of the other related regions - the river, lakes, and road.

Our representation and use of knowledge is similar to that described by Tenenbaum [5]. The main difference is in Tenenbaum's use of single pixel attributes to uniquely distinguish objects in a given context. We use the object attributes to aid in the segmentation of the image and then use the attributes of large segmented parts for recognition. Bajcsy and Lieberman [6] attempted to model general outdoor scenes, but the detailed models which we use are specific to the particular task being performed and scene being analyzed. Bajcsy and Tavakoli [7] used special purpose programs to locate specific objects (i.e. roads and bridges) rather than a general matching technique.

Matching Strategy

For simplicity, and because it proved to be accurate, we use a general purpose symbolic matching system to find the elements in the image which correspond to the individual elements in the user specified model [3]. The general matching procedure locates corresponding regions by comparing the symbolic description of a region in one image with the symbolic descriptions of regions in another image, and selecting the region which is most similar. The similarity function for two regions is essentially the sum of normalized absolute differences of the feature values of the two regions. This gives a region to region match function of:

$$R_{ij} = - \sum_{k=1}^n |v_{ik} - v_{jk}| w_k \quad (1)$$

where R_{ij} is the value of the match between regions i and j ,

n is the number of valid features, $V_{ik}(V_{jk})$ is the value of the k th feature of regions $i(j)$, and W_k is the normalization weight for the k th feature. Region i and j are from different images. The result is negative so that better matches are larger, i.e. closer to zero.

In the general image matching case, the user can specify that some features will change, or are less reliable for the given task. Those features are then given less weight in the computation of the similarity function. Conversely other features can be given as required matches and these are given more weight in the rating. These task-dependent weights for the feature value differences are in addition to the normalization factors which are constant for all tasks. The weights were selected so that a poor match in a highly weighted feature, i.e. a constant feature, will have the same impact as several poor matches in lower weighted features. The different weights should be less important when matching with the user specified model, since the given features are only the ones important for the description and thus are features which should not change.

The system locates the best corresponding segment in the image for each segment in the model, but no attempt is made to find the globally best one-to-one mapping. Because of this, the order in which corresponding segments are located may be important. Since we are using features such as neighboring regions, and regions above, below, left and right, an incorrect match early in the process will lead to inaccurate similarity ratings later. The default order for matching is by size, but other orders can be specified by the user or made available automatically if desired. We assume that the ordering information is supplied by the user, and have made no attempt to automate the strategy generation process as in [8].

We now present results of applying these procedures on two scenes.

Results

The first scene is the one presented in Figures 1 to 3 in the previous section. This is an aerial view of a mostly rural area. The goal, for this example, is to locate the docks along the lower edge of the large river. Figure 2 gives a pictorial description of the important elements in the scene, such as might appear on a map or a sketched description of the desired objects. The portion of the image where the docks can be located is specified in relation to the objects in the model description. Thus the bounds of this area can be computed when the position in the image of the objects in the model are known. This limited area can be further analyzed to complete the task. Figure 4 shows the lines and regions which are segmented in this image, because the lakes and river are the only regions used in the description. These are the only ones extracted by the segmentation operation.

When the user specified model is matched with the machine segmented image, the regions and line shown in Figure 5 are selected as the corresponding segments in the scene. The location and dimensions of the upper lake and the location of the lower edge of the river are used to specify the area where the docks are located. This area is extracted from the complex image for detailed analysis with which the individual docks may be located. The two subimages corresponding to the dock area are given in Figure 6.

The results for this scene require a few comments. The segmentation procedures failed to separate the lower lake



Figure 4. Segmentation of Image 1.



Figure 5. Regions Corresponding to Model Elements.



Figure 6. DOCKS Subimages.

from the river region since the two regions are connected by a narrow neck, but a separation should be possible with more processing of the extracted regions to eliminate narrow necks. In this case the neck was just wide enough to escape the processing which is already applied. Because of this, there is no actual region in the image which corresponds to the lower lake in the model. The region in the image which is selected as the corresponding region is the upper lake region. In the procedure which computes the position of the dock areas, the position of the best matching lake is used. Because of the errors in the relative position of the selected lower lake region, i.e. it is on the wrong side of the road, the upper lake was matched best and thus this is the region which is used. This illustrates the insensitivity of the higher level processing to certain errors in the lower level processing.

We also applied the same system to another scene, shown in Figure 7. This is an aerial photograph dominated by a river, a central power plant, a large group of houses and several prominent roads. The task for this scene is to locate the power plant area. This image has about 2000 x 2000 points so that a detailed analysis of the entire image would be impractical. The position of the power plant is specified by the location of certain roads and river segments. Figure 8 illustrates the important objects in the scene which we wish to locate. The location of the river regions and some of the roads is important mainly because of the relation of these segments to other, necessary, road segments. Figure 9 shows the segments extracted from this image. The segments include the large untextured regions, small bright regions, and long line features. Figure 10 shows the regions selected by the system as those which correspond to the elements in the user defined model. The position of the centrally located road segment is used to



Figure 7. Aerial Image 2.

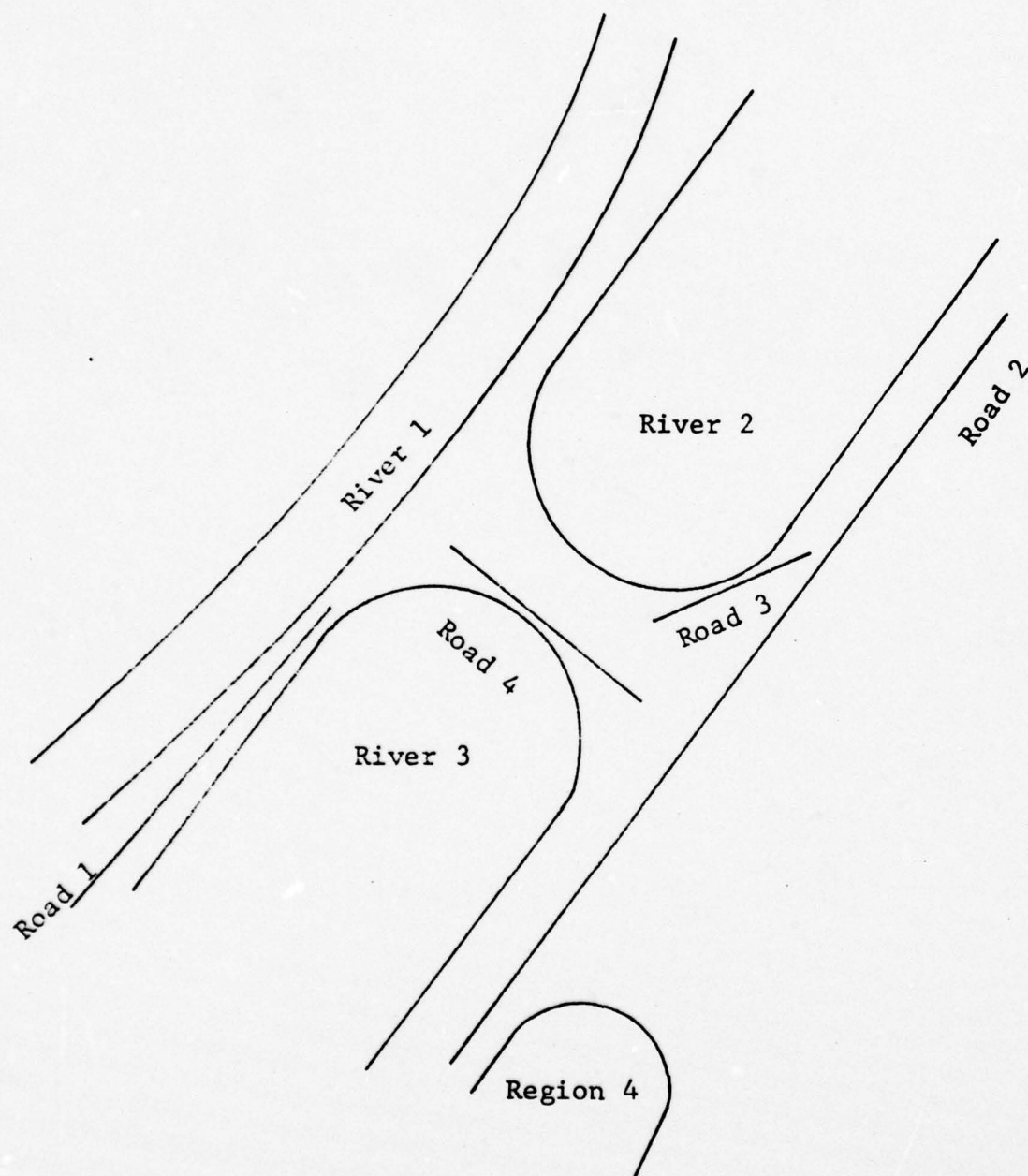


Figure 8. Pictorial Model of Power Plant Location.

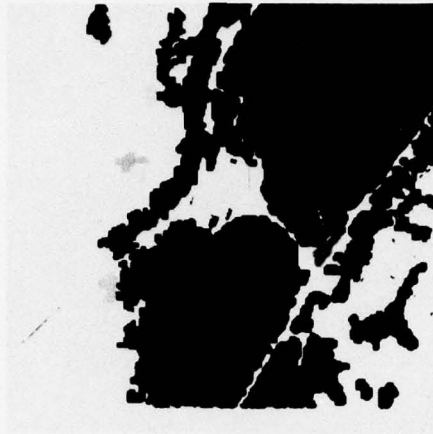


Figure 9. Segmentation of Image 2.

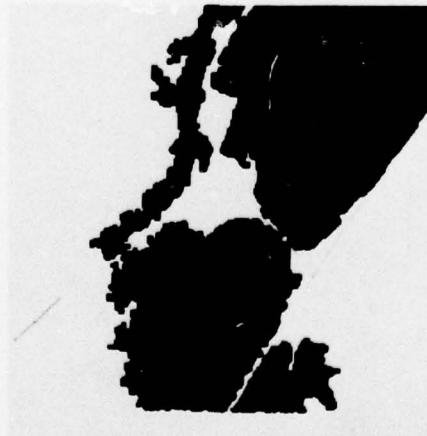


Figure 10. Regions Corresponding to the Model.

specify the subwindow which includes the power plant and very little else. Figure 11 shows the subwindow which was extracted from this image. This subwindow is 512 x 512 pixels and could be further analyzed by this system to find specific objects if a description of this portion of the scene is furnished by the user.

Other experiments were tried with this image and model to determine which features are necessary for locating the regions in the model. Correct results were obtained using only the size and relative positions (above, below, etc.) of the segments. This is accounted for mostly by the fact that the sizes of the regions are very large compared to other regions (the lines are also the longest lines which were extracted), and the sizes specified in the model were close to the sizes in the image. If the sizes in the model were less precise then other features would be necessary.

Conclusions

Some results of processing a complex aerial image using both the line and the region based techniques have been shown. It appears that the use of simple techniques, specifically suited to particular objects in an image, may allow useful processing of rather complex images. At this time the array of segmentation attributes is limited. While it is hoped that the described techniques have general applicability, our experience with real images is, as yet, limited.

References

1. R. Nevatia, "Locating Object Boundaries in Textured Environments," IEEE Transactions on Computers, Vol. 25, No. 11, November 1976, pp. 829-832.

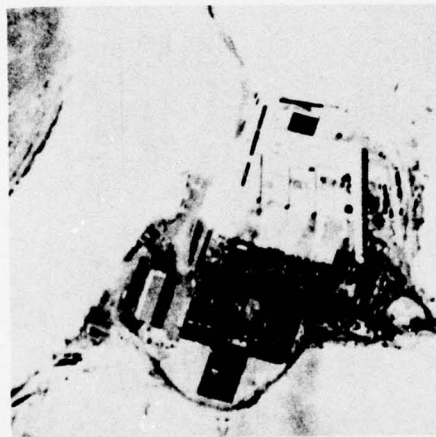


Figure 11. Extracted Power Plant Subimage.

2. R. Ohlander, "Analysis of Natural Scenes," Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania, April 1975.

3. K. Price, "Change Detection and Analysis in Multi-Spectral Images," Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, December 1976.

4. R. Nevatia and K. Price, "A Comparison of Some Segmentation Techniques," Proceedings of ARPA Image Understanding Workshop, April 1977, pp. 55-57.

5. J.M. Tenenbaum, "Locating Objects by Their Distinguishing Features in Multi-Sensory Images," Computer Graphics and Image Processing, Vol. 2, No. 3, December 1973.

6. R. Bajcsy and L.I. Lieberman, "Computer Description of Real Outdoor Scenes," Proceedings of the Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, August 1974, pp. 174-179.

7. R. Bajcsy and M. Tavakoli, "Computer Recognition of Roads from Satellite Pictures," IEEE Transactions on Systems, Man, and Cybernetics, SMC-6, 1976, pp. 623-637.

8. T.D. Garvey and J.M. Tenenbaum, "On the Automatic Generation of Programs for Locating Objects in Office Scenes," Proceedings of the Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, August 1974, pp. 162-168.

2.4 A New Edge Fitting Algorithm

Ikram E. Abdou

Introduction

In a previous paper [1] it was mentioned that the Hueckel operator is not an optimum edge fitting method because the operator is based on some approximations which affect the final performance. This paper introduces a new edge fitting algorithm which avoids some of the disadvantages of the Hueckel operator. A discussion of the basic ideas of the new algorithm and a comparison between its performance and the Hueckel performance are given in the following sections.

Basic Concepts

In edge fitting the image function $f(i,j)$ defined over a subregion \mathcal{A} is compared with an ideal edge model $S_{\omega}(i,j)$, where ω is a set of edge parameters. The difference between the actual and ideal models is function of ω , and by changing these parameters the difference can be minimized. The edge acceptance is based on the value of the minimum difference. If it is less than a given threshold (t), the image subregion is classified as an edge with the corresponding parameters ω_{\min} . Usually the mean square error is used to measure the difference between the ideal and actual edge. This error is given in the form

$$E_{\omega} = \sum_j \sum_i (S_{\omega}(i,j) - f(i,j))^2 \quad (1)$$

The ideal edge model used in the new algorithm is a central edge with slope along one of the four basic

directions (shown in Figure 1b). For this model ω is the set $\{\theta_i, n, a, \Delta\}$, where

- θ_i is the edge orientation
- n is the edge width
- a is the edge average height
- Δ is the slope of the central ramp.

E_ω is minimized by changing two discrete variables, θ_i and n , and two continuous variables, a and Δ . Since θ_i and n assume finite number of values, the minimization problem can be solved by repeating the computation for each discrete variable and choosing the minimum E_ω for all θ_i and n . As an example, assume that $f(i, j)$ is fitted to the vertical edge shown in Figure 1a, $S_\omega(i, j)$ is now given by

$$\begin{aligned} S_\omega(i, j) &= a - \Delta n & -N \leq i < -n, \forall j \\ &= a + \Delta i & -n \leq i \leq n, \forall j \\ &= a + \Delta n & n < i \leq N, \forall j \end{aligned} \quad (2)$$

Substituting in Eq. 1.

$$E_\omega = \sum_{i=-N}^{-(n+1)} \sum_j (a - \Delta n - f(i, j))^2 + \sum_{i=-n}^n \sum_j (a + \Delta i - f(i, j))^2 + \sum_{i=n+1}^N \sum_j (a + \Delta n - f(i, j))^2$$

Expanding and arranging terms. E_ω can be expressed as

$$\begin{aligned} E_\omega = & \sum_i \sum_j (a - f(i, j))^2 + \Delta \left\{ 2n \sum_{i=-N}^{-(n+1)} \sum_j f(i, j) - 2 \sum_{i=-n}^n \sum_j i f(i, j) \right. \\ & \left. - 2n \sum_{i=n+1}^N \sum_j f(i, j) \right\} + \Delta^2 \left\{ \sum_{i=-N}^{-(n+1)} \sum_j n^2 + \sum_{i=-n}^n \sum_j i^2 + \sum_{i=n+1}^N \sum_j n^2 \right\} \end{aligned} \quad (3.a)$$

$$= C_0 + C_1 \Delta + C_2 \Delta^2 \quad (3.b)$$

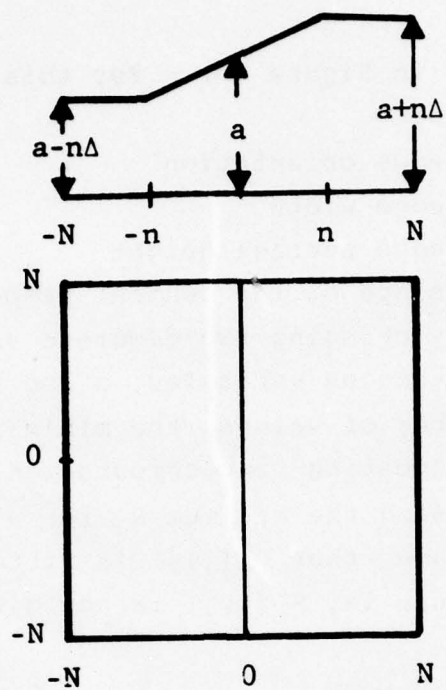


Figure 1a. $S_{\omega}(i,j)$ for Ideal Vertical Edge.

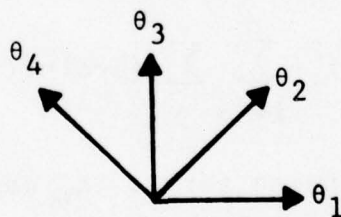


Figure 1b. The Basic Four Directions.

For given edge orientation and width, E_ω is minimum when

$$\frac{\partial E_\omega}{\partial a} = 0 \quad (4.a)$$

$$\frac{\partial E_\omega}{\partial \Delta} = 0 \quad (4.b)$$

Substituting in Eq. (3)

$$a = \frac{1}{(2N+1)^2} \sum_i \sum_j f(i,j) \quad (5.a)$$

$$\Delta = -\frac{C_1}{2C_2} \quad (5.b)$$

where C_1 and C_2 are as given in Eq. 3. In addition

$$E_{\min} = C_0 - \frac{C_1^2}{4C_2} \quad (6)$$

Similar expressions can be obtained for different edge orientation and widths. The actual image is fitted to the edge model which results in minimum error. In the evaluation of Eq. 6 it is not necessary to calculate C_0 because it is constant for the same image subregion. The acceptance of the edge can be based on the values of Δ_{\min} and E_{\min} . The computation time required by the previous procedure is much less than the Hueckel time. While the results obtained are more accurate as will be shown in the following section.

Experimental Results

The performance of the edge fitting algorithm with different mask sizes is evaluated using the figure of merit defined in (2). The results obtained are shown in Figure 2. Comparing these results with the results obtained for the Hueckel operator (1), it is clear that the new algorithm has better performance.

Another question of interest is to what extent is the new edge fitting operator better than a 3-level simple operator (3) of the same mask size? The performance curves of 3-level simple operator with mask sizes 5, and 9 are shown in Figure 2. It is clear that for small mask size and very low SNR the edge fitting algorithm is not so good as simple masks operator. This observation can be explained by the fact that the edge fitting algorithm bases its decision on an estimation of the edge parameters. This estimation is sensitive to noise specially when the number of pixels used is small. However the edge fitting operator has better performance for high SNR and for large mask sizes.

References

1. I.E. Abdou and W.K. Pratt, "Some Comments on the Hueckel Operator," University of Southern California, Image Processing Institute, USCIPi Report 770, September 1977, pp. 30-36.
2. W.K. Pratt, Digital Image Processing, Wiley-Interscience, New York, 1978.
3. G.S. Robinson, "Detection and Coding of Edge Using Directional Masks," University of Southern California, Image Processing Institute, USCIPi Report 660, March 1976, pp. 40-57.

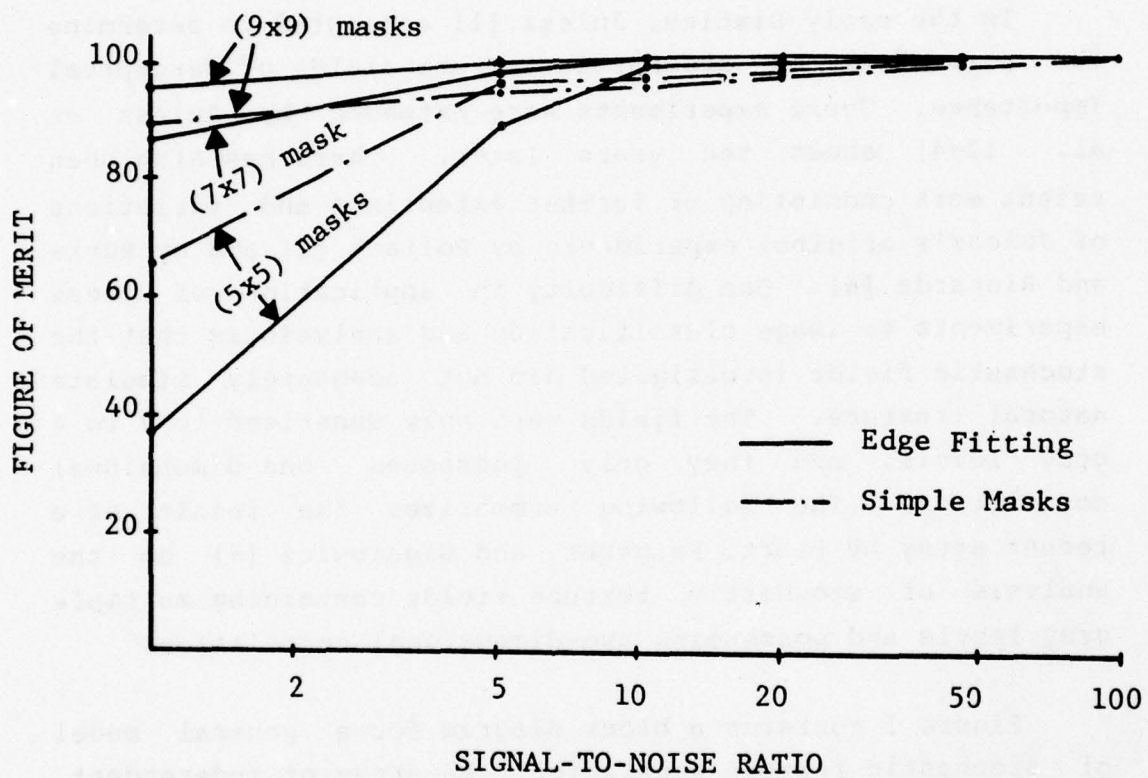


Figure 2. Edge location figure of Merit as a function of signal-to-noise ratio.

2.5 Stochastic Texture Analysis

William K. Pratt*

In the early Sixties, Julesz [1] attempted to determine the parameters of stochastic texture fields of perceptual importance. These experiments were extended by Julesz et al. [2-4] about ten years later. There has also been recent work consisting of further extensions and variations of Julesz's original experiments by Pollack [5] and by Purks and Richards [6]. One difficulty in application of these experiments to image classification and analysis is that the stochastic fields investigated did not adequately simulate natural texture. The fields were only quantized to 2 to 4 gray levels, and they only possessed one-dimensional correlation. The following summarizes the results of a recent study by Pratt, Faugeras, and Gagalowicz [7] on the analysis of stochastic texture fields containing multiple gray levels and possessing two-dimensional correlation.

Figure 1 contains a block diagram for a general model of stochastic texture generation. An array of independent, identically distributed pixels $W(j,k)$ with probability density $p(W)$ passes through a linear or nonlinear system described by the operation $\mathcal{O}\{\cdot\}$ to produce the stochastic texture field $F(j,k)$. Reference [7] provides design methods for selecting $p(W)$ and $\mathcal{O}\{\cdot\}$ in order to achieve textural fields with desired statistical properties.

*This work was performed in collaboration with O.D. Faugeras and A. Gagalowicz of Institut de Recherche d'Informatique et d'Automatique, Le Chesnay, France.

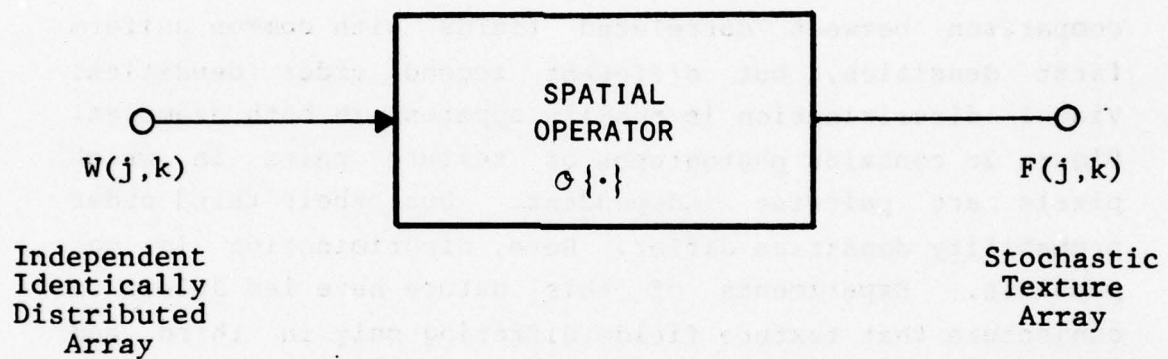
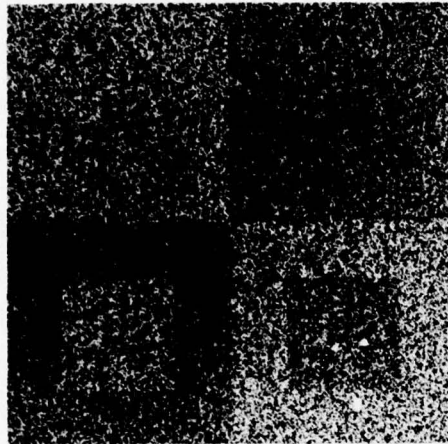


Figure 1. Stochastic Texture Generation Model.

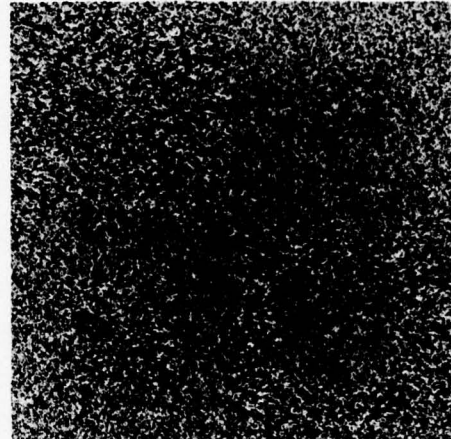
Figure 2 contains computer-generated texture fields for an experiment to determine perceptual sensitivity to differences in probability density between texture field pairs. Moments listed on the figure are defined elsewhere. Figure 2a offers a comparison of 256 x 256 pixel fields of 256 gray levels with different independent first order densities but identical means. The A field possesses a uniform density, while the B field has a density shaped like an isosceles triangle. In figure 2b, there is a split field comparison between correlated fields with common uniform first densities, but different second order densities. Visual discrimination is readily apparent in both examples. Figure 2c contains photographs of texture pairs in which pixels are pairwise independent. But, their third order probability densities differ. Here, discrimination is not possible. Experiments of this nature have led Julesz to conjecture that texture fields differing only in third and higher order densities cannot be discriminated by a human observer. One possible explanation for the lack of discrimination in figure 2c is the absence of spatial correlation. However, this is not the proper explanation. In figure 2d the random numbers of figure 2c have been spatially correlated by separable linear processing along rows and columns to produce fields with first order Markovian correlation, but differing third order densities. Since discrimination is not possible, Julesz's conjecture has been confirmed for spatially correlated as well as uncorrelated fields.

The experiments of figure 3 illustrate visual sensitivity to changes in the autocorrelation of Gaussian density texture fields with identical first order densities. Texture differences corresponding to changes of about 10% in the correlation lag parameters are visible.



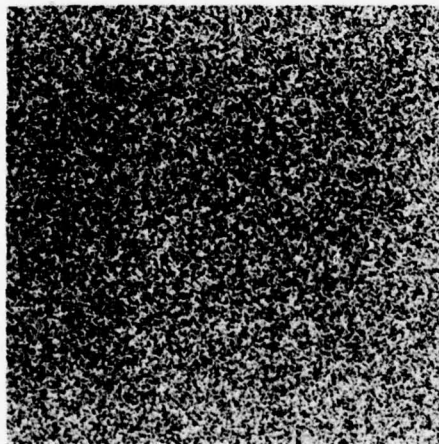
(a) Uncorrelated first order density fields

$$\sigma_A = 0.289 \quad \sigma_B = 0.204$$



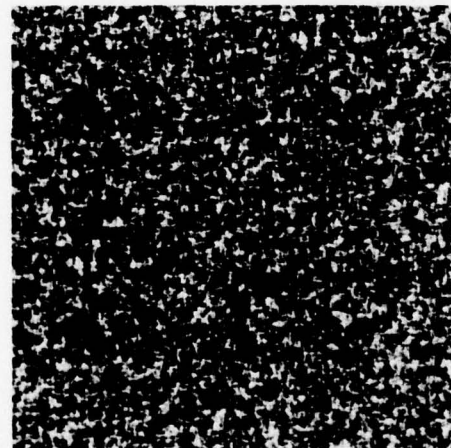
(b) Correlated second order density fields

$$\begin{aligned} \sigma_A &= 0.289 & \sigma_B &= 0.289 \\ \alpha_A &= 0.250 & \alpha_B &= -0.250 \end{aligned}$$



(c) Uncorrelated third order density fields

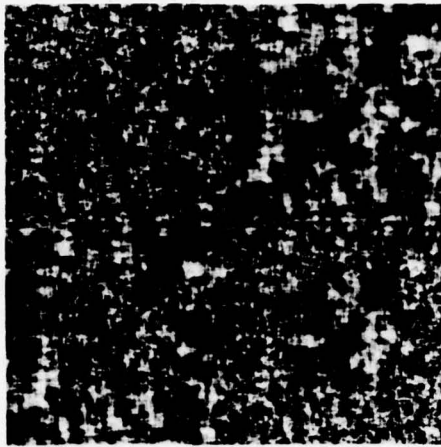
$$\begin{aligned} \sigma_A &= 0.289 & \sigma_B &= 0.289 \\ \alpha_A &= 0.000 & \alpha_B &= 0.000 \\ \theta_A &= 0.058 & \theta_B &= -0.058 \end{aligned}$$



(d) Correlated third order density fields

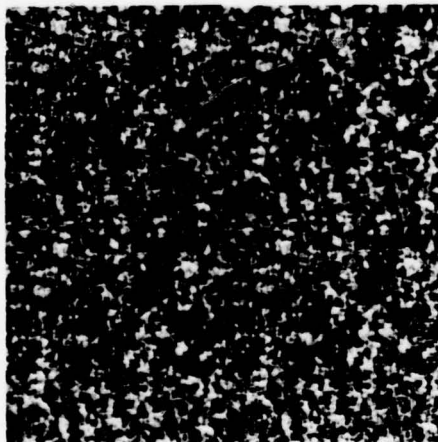
$$\begin{aligned} \sigma_A &= 0.167 & \sigma_B &= 0.167 \\ \alpha_A &= 0.750 & \alpha_B &= 0.750 \\ \theta_A &= -0.057 & \theta_B &= 0.077 \end{aligned}$$

Figure 2. Field comparison of texture fields with differing probability densities. $\eta_A = \eta_B = 0.500$.



(a) No common lags

$$\begin{aligned}\sigma_A &= 0.167 & \sigma_B &= 0.167 \\ \alpha_A &= 0.750 & \alpha_B &= 0.900 \\ \beta_A &= 0.563 & \beta_B &= 0.810 \\ \gamma_A &= 0.422 & \gamma_B &= 0.729\end{aligned}$$



(b) First lag common

$$\begin{aligned}\sigma_A &= 0.167 & \sigma_B &= 0.167 \\ \alpha_A &= 0.750 & \alpha_B &= 0.750 \\ \beta_A &= 0.563 & \beta_B &= 0.700 \\ \gamma_A &= 0.422 & \gamma_B &= 0.490\end{aligned}$$



(c) First and second lags common

$$\begin{aligned}\sigma_A &= 0.167 & \sigma_B &= 0.167 \\ \alpha_A &= 0.750 & \alpha_B &= 0.750 \\ \beta_A &= 0.563 & \beta_B &= 0.563 \\ \gamma_A &= 0.422 & \gamma_B &= 0.550\end{aligned}$$

Figure 3. Field comparison of Gaussian density texture fields.
 $\eta_A = \eta_B = 0.500$.

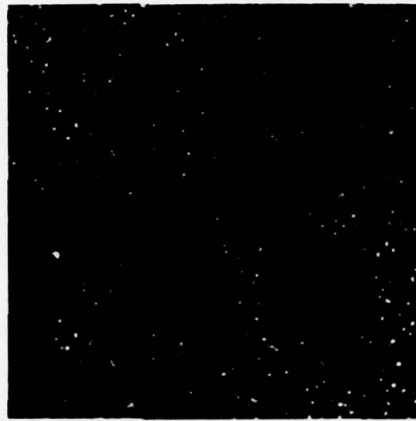
Figure 4 proves the insufficiency of the autocorrelation function as a description of texture. In figures 4a and 4b the fields are generated with identical means, variances, and autocorrelation functions, but different third order densities. The texture differences are easily discernible.

The experiments of figures 2 to 4 serve to establish useful bounds for developing stochastic-based visual texture features. Second order statistical measures should be sufficient, but the mean, variance, and autocorrelation function measures, by themselves, although directly or indirectly necessary, are not sufficient. The task then is to determine those stochastic features within the established bounds that are perceptually sufficient.

The stochastic texture examples presented here are important not only for the inherent information they convey about visual texture perception; they also provide valuable test fields for evaluation of candidate texture features. A universally sufficient texture feature vector must provide sufficient information to quantitatively discriminate the Gaussian texture fields of figure 3 possessing differing autocorrelation functions, and also discriminate the non-Gaussian fields of figure 4 that have identical autocorrelation. If such feature vectors pass this stringent test, then they are worth evaluating for natural texture.

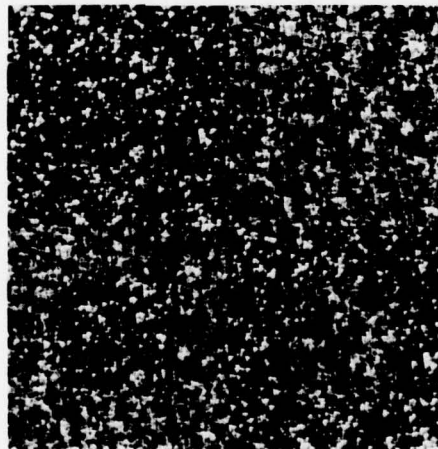
References

1. B. Julesz, "Visual Pattern Discrimination," IRE Transactions on Information Theory, Vol. IT-8, No. 1, February 1962, pp. 84-92.



(a) fields with same random input,
but different spatial operators
modified uniform density input
(uniform numbers raised to 256th
power)

$\eta_A = 0.413$	$\eta_B = 0.412$
$\sigma_A = 0.078$	$\sigma_B = 0.078$
$\alpha_A = 0.915$	$\alpha_B = 0.917$
$\theta_A = 1.512$	$\theta_B = 0.006$



(b) fields with different random inputs,
but same spatial operators
field A - modified uniform density input
field B - uniform density input

$\eta_A = 0.500$	$\eta_B = 0.500$
$\sigma_A = 0.167$	$\sigma_B = 0.167$
$\alpha_A = 0.750$	$\alpha_B = 0.750$
$\theta_A = 0.840$	$\theta_B = -0.010$

Figure 4. Field comparison of texture fields with common mean, variance, and autocorrelation function, but different densities.

2. B. Julesz, et al., "Inability of Humans to Discriminate Between Visual Textures that Agree in Second-Order Statistics - Revisited," Perception, Vol. 2, 1973, pp. 391-405.
3. B. Julesz, Foundations of Cyclopean Perception, University of Chicago Press, 1971.
4. B. Julesz, "Experiments in the Visual Perception of Texture," Scientific American, Vol. 232, No. 4, April 1975, pp. 2-11.
5. I. Pollack, Perceptual Psychophysics, Vol. 13, 1973, pp. 276-280.
6. S.R. Purks and W. Richards, "Visual Texture Discrimination Using Random-Dot Patterns," Journal of the Optical Society of America, Vol. 67, No. 6, June 1977, pp. 765-771.
7. W.K. Pratt, O.D. Faugeras, and A. Gagalowicz, "Visual Discrimination of Stochastic Texture Fields," February 1978, submitted for publication.

2.6 Singular Value Decomposition Image Feature Extraction

Behnam Ashjari and William K. Pratt

The singular value decomposition (SVD) is a numerical technique of matrix transformation by which an arbitrary matrix can be expressed in outer product form. SVD expansions have been applied in the solutions of ill-conditioned sets of linear equations [1]. There has

also been recent application of the SVD concept to image restoration and coding [2,3]. Another application, explored here, is the use of the SVD to extract descriptive features from an image region for purposes of classification or analysis.

Singular Value Matrix Decomposition

Consider an $N \times N$ matrix \underline{F} of rank R containing non-negative, real elements. The SVD transform of \underline{F} is defined to be

$$\underline{S} \equiv \underline{\Lambda}^{\frac{1}{2}} = \underline{U}^T \underline{F} \underline{V} \quad (1)$$

where \underline{U} and \underline{V} are unitary matrices and \underline{S} is a diagonal matrix whose diagonal elements $\lambda^{\frac{1}{2}}(1) \geq \lambda^{\frac{1}{2}}(2) \geq \dots \geq \lambda^{\frac{1}{2}}(R) = \dots \lambda(N)$ are called the singular values of \underline{F} . Since \underline{U} and \underline{V} are unitary matrices, the inverse transform yields

$$\underline{F} = \underline{U} \underline{S} \underline{V}^T \quad (2)$$

The unitary matrix \underline{U} contains columns \underline{u}_n composed of eigenvectors of the symmetric matrix product $\underline{F} \underline{F}^T$. Similarly, the columns \underline{v}_n of \underline{V} are eigenvectors of $\underline{F}^T \underline{F}$. The defining relations are

$$\underline{S} \underline{S}^T = \underline{U}^T [\underline{F} \underline{F}^T] \underline{U} = \underline{\Lambda} \quad (3a)$$

$$\underline{S}^T \underline{S} = \underline{V}^T [\underline{F}^T \underline{F}] \underline{V} = \underline{\Lambda} \quad (3b)$$

Thus, the squares of the singular values arranged in the diagonal matrix $\underline{\Lambda}$ are the eigenvalues of both $\underline{F} \underline{F}^T$ and $\underline{F}^T \underline{F}$. Consequently, the singular values can be algebraically generated. From solution of the N roots of the characteristic equation

$$C(\lambda) \equiv \sum_{j=0}^N a_j \lambda^j = \det[\underline{F}\underline{F}^T - \lambda \underline{I}] \quad (4)$$

It should be observed that the eigenvalues $\underline{\lambda}$ are positive, real variables since $\underline{F}\underline{F}^T$ and $\underline{F}^T\underline{F}$ are real, symmetric matrices. For consistency, the singular values \underline{s} are chosen to be the positive square roots of the corresponding eigenvalues.

It is often useful to express the matrix decomposition of eq.(2) in the outer product expansion

$$\underline{F} = \sum_{j=1}^R S(j) \underline{u}_j \underline{v}_j^T \quad (5)$$

where the sum is over the rank R of \underline{F} . In a similar manner, the symmetric matrix products can be expressed in the outer product forms

$$\underline{F}\underline{F}^T = \sum_{m=1}^N \underline{f}_m \underline{f}_m^T \quad (6a)$$

$$\underline{F}^T\underline{F} = \sum_{m=1}^N \underline{h}_m \underline{h}_m^T \quad (6b)$$

where \underline{f}_m denotes the m -th column of \underline{F} and \underline{h}_m is the m -th row of \underline{F} .

SVD Image Feature Extraction Concept

The singular values of a matrix can be considered as descriptors or features of the matrix elements and their inter-relationships. If the matrix is composed of randomly chosen real numbers, the singular values will tend toward equality. On the other hand, a highly structured matrix will exhibit a few dominating singular values. Figure 1 sketches the qualitative structural relationship of the singular values. This observation forms the basis for

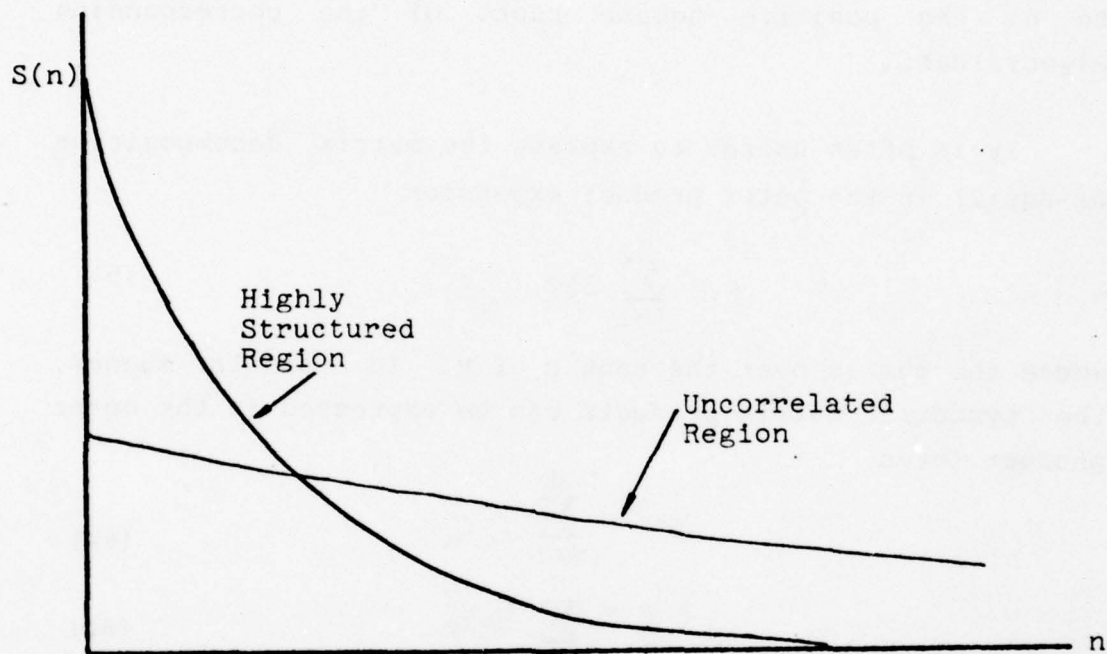


Figure 1. Example of Singular Value Behavior of Image Regions.

utilization of the SVD as a means of forming image features. An SVD expansion is formed over a block of $N \times N$ pixels, and some measure of the skewness of the singular value distribution is formed to characterize the spatial "coherence" of the pixel values. This vague concept will be solidified shortly.

Deterministic Properties

Certain deterministic image patterns possess SVD expansions that can be easily computed. For example, an array of unit value pixels can be generated by the outer product expansion

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} [1 \ 1 \ . \ . \ . \ 1] = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 1 & 1 & . & . & . & 1 \\ \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & & & & \vdots \\ 1 & 1 & . & . & . & 1 \end{bmatrix} \quad (7)$$

Such a matrix is of unit rank, and therefore, possesses only one non-zero singular value. A matrix containing alternating vertical stripes of zeroes can be formed by

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} [1 \ 0 \ 1 \ 0 \ . \ . \ . \ 1 \ 0] = \begin{bmatrix} 1 & 0 & 1 & 0 & . & . & . & 1 & 0 \\ 1 & 0 & 1 & 0 & . & . & . & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots & \vdots \\ 1 & 0 & 1 & 0 & . & . & . & 1 & 0 \end{bmatrix} \quad (8)$$

Horizontally striped arrays can be generated by reversing the positions of the vectors in eq.(8). The reason that the arrays of eqs.(7) and (8) have only one non-zero singular value is that their rows are linearly dependent, in fact identical. Hence, any array with repeated rows (or columns)

will possess only one non-zero singular value. This class includes striped arrays of various periods and arrays with single horizontal or vertical line strokes.

A checkerboard array of ones and zeroes can be generated by the sum of two outer products

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ 1 \\ 0 \end{bmatrix} [1010\dots10] + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} [0101\dots01] = \begin{bmatrix} 1010\dots10 \\ 0101\dots01 \\ \vdots \\ \vdots \\ 1010\dots10 \\ 0101\dots01 \end{bmatrix} \quad (9)$$

and therefore, possesses two non-zero singular values. Finally, an $N \times N$ identity matrix with ones down its main diagonal and zeroes elsewhere is a rank N matrix and has N equal singular values.

These simple deterministic examples illustrate some interesting points. First, the SVD expansion is invariant to the period of striped horizontal and vertical patterns. In fact, a constant amplitude array, which can be considered an array with a 100% duty cycle period, has the same number of non-zero singular values (one) as a striped matrix with single pixel stripes. As a consequence, the SVD is clearly not useful as a measure of periodic structure of deterministic arrays. The second major point of the examples is that rotation of a pattern affects the number of singular values. A checkerboard pattern consisting of diagonal stripes possesses two non-zero singular values, while horizontally or vertically striped arrays only have one non-zero singular value. As an even more extreme example consider the $N \times N$ diagonal matrix with a single diagonal line. It has N identical singular values. But, an array with a single pixel vertical line possesses one

non-zero singular value.

An interesting and useful deterministic property of the SVD is its invariance to unitary transforms. Consider a unitary transformation

$$\underline{Z} = \underline{L}_C^T \underline{F} \underline{L}_R \quad (10)$$

where $\underline{L}_C \underline{L}_C^{*T} = \underline{L}_R \underline{L}_R^{*T} = \underline{I}$. It is easily demonstrated that the characteristic equations of $\underline{Z} \underline{Z}^T$ and $\underline{F} \underline{F}^T$ are identical, and hence, the singular values of \underline{F} and \underline{Z} are the same. The eigenvector matrices are related by

$$\underline{U}_Z = \underline{L}_C^T \underline{U}_F \quad (11a)$$

$$\underline{V}_Z = \underline{L}_R^T \underline{V}_F \quad (11b)$$

As a consequence of this property, it is observed that the singular values of an image block and its two-dimensional Fourier or Hadamard transform, for example, are identical. Certainly, the deterministic structure of such arrays is substantially different.

The conclusion of these observations is that the singular value distribution is not well suited for the characterization of deterministic line, spot, or shape structure. But rather, the SVD does provide an indication of the structural dependence between rows and columns of a pixel array. This property, as will be shown, is extremely important for the characterization of texture-like regions of an image.

Stochastic Properties

If an image block \underline{F} is considered to be a sample of a two-dimensional random process, then the singular values $s(n)$, as defined by eq.(1), will be random variables since

the SVD is a linear transformation of \underline{F} to \underline{S} . Let \underline{F} be column scanned to produce the $N^2 \times 1$ vector \underline{f} [4]. Then, the image block may be stochastically defined by its joint probability density $f(\underline{f})$. It is also useful to describe the stochastic nature of the image block in terms of its mean vector $\underline{\eta}_f$ and its covariance matrix \underline{K}_f . With this information available, the joint density of the singular values $P[S(1), S(2), \dots, S(N)]$ and the singular value moments can be found, in theory, however, in practice, derivation is difficult.

Consider first, calculation of the moments of \underline{S} . Referring to eqs.(2) and (4), the mean matrix of the image block can be written as

$$E\{\underline{F}\} \equiv \underline{M}_F = E\{\underline{USV}^T\} = \sum_{n=1}^R E\{S(n) \underline{u}_n \underline{v}_n^T\} \quad (12)$$

Similarly, from eq.(3), the means of the symmetric matrix products are

$$E\{\underline{FF}^T\} = E\{\underline{UAU}^T\} \quad (13a)$$

$$E\{\underline{F}^T \underline{F}\} = E\{\underline{VAV}^T\} \quad (13b)$$

with reference to eq.(6), the left-hand terms of eq.(13) can be expressed as

$$E\{\underline{FF}^T\} = \sum_{n=1}^N E\{\underline{f}_n \underline{f}_n^T\} = \sum_{n=1}^N \underline{R}_{Cn} \quad (14a)$$

$$E\{\underline{F}^T \underline{F}\} = \sum_{n=1}^N E\{\underline{h}_n \underline{h}_n^T\} = \sum_{n=1}^N \underline{R}_{Rn} \quad (14b)$$

where \underline{R}_{Cn} and \underline{R}_{Rn} denote the n -th column and row correlation matrices of \underline{F} , respectively. Continuing one step further, the means of the symmetric matrix products about the mean of \underline{F} are given by

$$E\{(\underline{F}-\underline{M}_F)(\underline{F}-\underline{M}_F)^T\} = \sum_{n=1}^N \underline{K}_{Cn} \quad (15a)$$

$$E\{(\underline{F}-\underline{M}_F)^T(\underline{F}-\underline{M}_F)\} = \sum_{n=1}^N \underline{K}_{Rn} \quad (15b)$$

where \underline{K}_{Cn} and \underline{K}_{Rn} are column and row covariance matrices of \underline{F} .

The preceding equations provide a link

$$E\{\underline{U}\underline{A}\underline{U}^T\} = \sum_{n=1}^N \underline{R}_{Cn} = \sum_{n=1}^N \underline{K}_{Cn} + \underline{M}_F \underline{M}_F^T \quad (16a)$$

$$E\{\underline{V}\underline{A}\underline{V}^T\} = \sum_{n=1}^N \underline{R}_{Rn} = \sum_{n=1}^N \underline{K}_{Rn} + \underline{M}_F^T \underline{M}_F \quad (16b)$$

between the mean values of the eigen-expansions of \underline{F} and its first and second order moments. A special case of practical interest occurs when the mean matrix \underline{M}_F is constant valued and the row and column covariance matrices are identical from row-to-row and column-to-column. For this case,

$$E\{\underline{U}\underline{A}\underline{U}^T\} = N \underline{K}_C + \underline{M}_F \underline{M}_F^T \quad (17a)$$

$$E\{\underline{V}\underline{A}\underline{V}^T\} = N \underline{K}_R + \underline{M}_F^T \underline{M}_F \quad (17b)$$

where \underline{K}_C and \underline{K}_R denote the column and row covariance matrices, respectively. Thus, the mean eigen-expansions assume the same forms as the array column and row correlation matrices. Unfortunately, it is not possible to proceed further with this approach and isolate the mean of $\underline{\Lambda}$ (equivalently, the second moment of \underline{S}) from eqs.(16) or (17) since \underline{U} and \underline{V} are dependent upon $\underline{\Lambda}$.

The moments of \underline{S} can be found, in principle, by direct computation of the means of the eigenvalues of the characteristic equation

$$C(\lambda) = \det[\underline{F}\underline{F}^T - \lambda \underline{I}] = \sum_{i=0}^N C_i \lambda^i \quad (18)$$

where C_i is a function of the elements of \underline{F} . The singular values of \underline{F} are the positive square roots of the eigenvalue roots of the characteristic equation. Isolation of the roots of eq.(17) in terms of \underline{F} and subsequent moment calculation has proved complex even for general 2 x 2 pixel arrays. Extension of the approach to large size arrays seems to be totally intractable. However, there is some benefit to be derived from consideration of the characteristic equation from a stochastic viewpoint.

Consider the case in which \underline{F} is a sample from an array of correlated wide sense stationary random variables with common variance σ^2 and a separable covariance matrix $\underline{K} = \underline{K}_C \otimes \underline{K}_R$, where \underline{K}_C and \underline{K}_R represent column and row covariance matrices, respectively. Now, let a Karhunen-Loeve transform be performed on columns and rows of \underline{F} to produce the random array

$$\underline{Z} = \underline{L}_C^T \underline{F} \underline{L}_R \quad (19)$$

where

$$\underline{L}_C^T \underline{K}_C \underline{L}_C = \underline{\Gamma}_C \quad (20a)$$

$$\underline{L}_R^T \underline{K}_R \underline{L}_R = \underline{\Gamma}_R \quad (20b)$$

and \underline{L}_C , \underline{L}_R , $\underline{\Gamma}_C$, $\underline{\Gamma}_R$ are eigenvector and eigenvalue matrices of \underline{K}_C and \underline{K}_R , as noted. From eq.(11), it has been established that the singular values of a matrix are invariant to a two-dimensional unitary transformation because the characteristic equation is not changed by such a transformation. Since the K-L transform is unitary, it is concluded that the singular values of the correlated array \underline{F} and the uncorrelated array \underline{Z} are identical. This may seem

to indicate that the SVD is useless for characterizing the structure of correlated arrays. But, this is not so. The spatial correlation information of \underline{F} has simply been converted to another form - the variance distribution of \underline{Z} . To show this, let \underline{V}_Z be an $N \times N$ matrix of the variance terms of \underline{Z} in spatial correspondence with \underline{Z} . Then, it is easily shown that

$$\underline{V}_Z = \underline{Y}_C \underline{Y}_R^T \quad (21)$$

where the vectors \underline{Y}_C and \underline{Y}_R contain the diagonal terms of $\underline{\Gamma}_C$ and $\underline{\Gamma}_R$, respectively. For some well-defined processes, such as the first order Markov process, \underline{V}_Z can be determined analytically. Thus, to summarize, if \underline{F} is wide sense stationary with a separable covariance matrix, then the moments of \underline{F} can be found indirectly from the moments of \underline{Z} which is an uncorrelated array with an energy distribution given by eq.(21).

Attention will now be directed toward the establishment of the joint probability density of the singular values of \underline{F} , and from this, the singular value moments. Consider the outer product expansion

$$\underline{G} = (\underline{F} - \underline{M}_F)(\underline{F} - \underline{M}_F)^T = \sum_{n=1}^N (\underline{f}_n - \underline{\eta}_n)(\underline{f}_n - \underline{\eta}_n)^T \quad (22)$$

of \underline{F} about its mean \underline{M}_F where \underline{f}_n denotes the n -th column of \underline{F} and $\underline{\eta}_n$ is its mean (the n -th column of \underline{M}_F). Suppose that \underline{F} is considered as a collection of N independently generated N -dimensional vectors \underline{f}_n each with mean $\underline{\eta}_n$ but common covariance matrix \underline{K}_C . Then, \underline{G} is equivalent to the well known scatter matrix of multivariate statistical analysis [5,p.547]. The probability density of the scatter matrix has been established for Gaussian random vectors. In this case

$$P(\underline{G}) = W(\underline{G}, N, N, \underline{K}_C^{-1}) \quad (23)$$

where

$$W(\underline{G}, K, N, \underline{K}_C^{-1}) \equiv \frac{|\underline{K}_C^{-1}|^{\frac{N}{2}} |\underline{G}|^{\frac{1}{2}(N-K-1)} \exp\left\{-\frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \underline{K}_C^{-1}(i,j) G(i,j)\right\}}{\frac{K N}{2} \frac{K(K-1)}{4} \pi \Gamma\left(\frac{N}{2}\right) \Gamma\left(\frac{N-1}{2}\right) \cdots \Gamma\left(\frac{N-K+1}{2}\right)} \quad (24)$$

is called the Wishart density [6] and $\Gamma(\cdot)$ is the gamma function.

The next step is to determine the joint probability density of the eigenvalues of \underline{G} . In the late thirties several mathematicians found this density for a scatter matrix defined about its experimental mean with $\underline{K}_C = \sigma^2 \mathbf{I}$ [5, p.568]. From these results the eigenvalues of \underline{G} are found to possess a distribution with an element probability

$$P(\underline{\Lambda}) d\lambda(1) \dots d\lambda(N) = \frac{\prod_{i=1}^N \left[\prod_{j=1}^N \lambda(i) \right]^{-\frac{1}{2}} \prod_{i>j=1}^N [\lambda(j) - \lambda(i)] \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^N \lambda(i)\right\}}{(2\sigma^2)^{\frac{N^2}{2}} \prod_{i=1}^N \left[\Gamma\left(\frac{N+1-i}{2}\right) \right]^2} d\lambda(1) \dots d\lambda(N) \quad (25)$$

for $0 < \lambda(n) < \dots < \lambda(1) < \infty$. The second moment of the n -th singular value can then be found by direct evaluation of the integral

$$E\{S^2(n)\} = \int_0^\infty \int_0^{\lambda(1)} \dots \int_0^{\lambda(N-1)} \lambda(n) P(\underline{\Lambda}) d\lambda(N) \dots d\lambda(2) d\lambda(1) \quad (26)$$

No simple closed form solution for the general form of this integral equation has been found so far. But for the 2×2 case, the closed form solution has been derived as the

following:

$$\begin{aligned} E\{S^2(1)\} &= \sigma^2(2 + \frac{\pi}{2}) \\ E\{S^2(2)\} &= \sigma^2(2 - \frac{\pi}{2}) \end{aligned} \quad (27)$$

In 1960 James [7] generalized the solution for the distribution of eigenvalues of a scatter matrix for an arbitrary column population covariance matrix \underline{K}_C . The result is a complicated functional of zonal polynomials. Nevertheless, the distribution can be numerically evaluated, and its second moments can be determined in a similar fashion to eq.(26). Then by the diagonalization procedure of eqs.(19) to (21), it is possible to relate the second moment of the singular values to the degree of spatial correlation of \underline{F} for a wide sense stationary process with a separable covariance function.

Figure 2 contains a plot of the singular value variance as a function of the spatial correlation factor ρ for a zero mean, unit variance Gaussian process with a separable covariance matrix $\underline{K} = \underline{K}_C \otimes \underline{K}_R$ where $\underline{K}_C = \underline{K}_R = [\rho^{|i-j|}]$. These theoretical results support the previous contention that the singular value distribution tends toward uniformity as ρ approaches zero, and toward skewness for ρ large.

Experimental Results

Figure 3 contains two images of natural texture, grass and ivy, along with two artificially generated fields obtained by convolutional processing of two-dimensional fields of randomly generated pixels. The subjective match between the natural and artificial fields appears to be in reasonable agreement. Figures 4 and 5 contain plots of the singular values extracted from 16 x 16 pixel blocks in the center of each image. The distributions of the natural

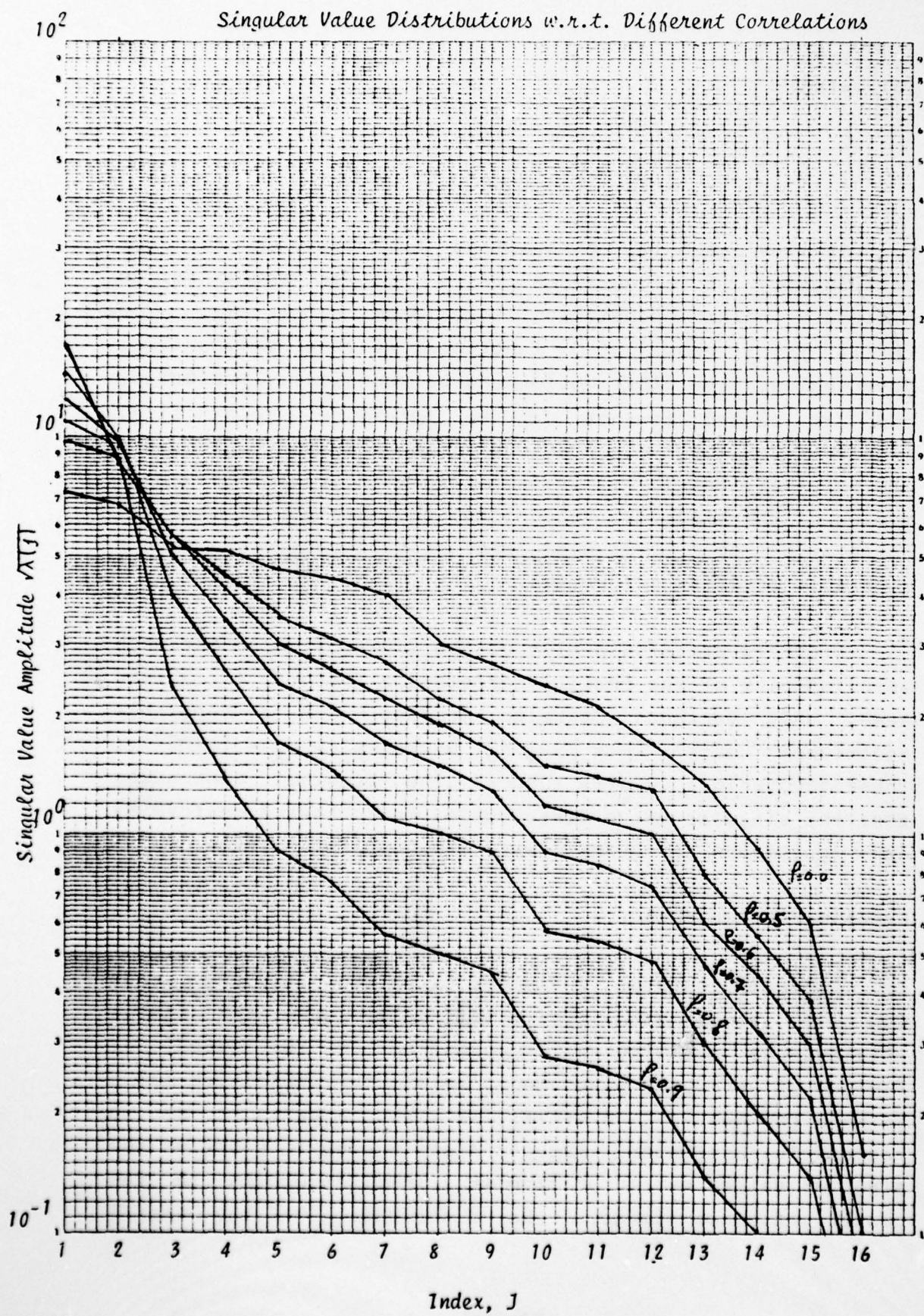
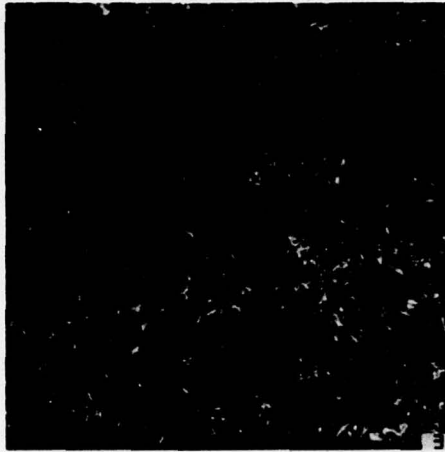
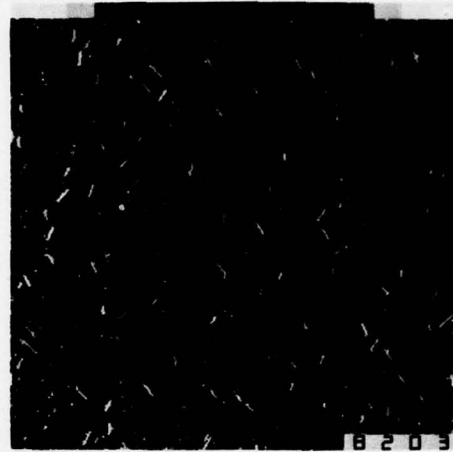


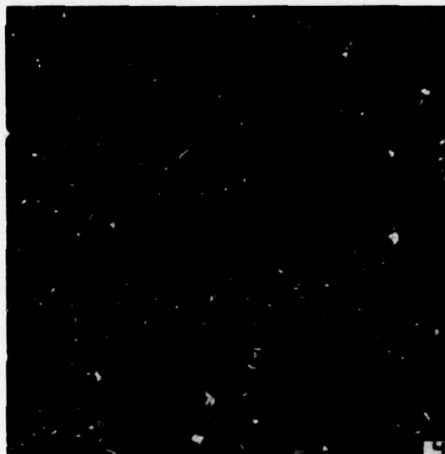
Figure 2.



(a) natural grass



(b) artificial grass



(c) natural ivy



(d) artificial ivy

Figure 3. Examples of Natural and Artificial Texture.

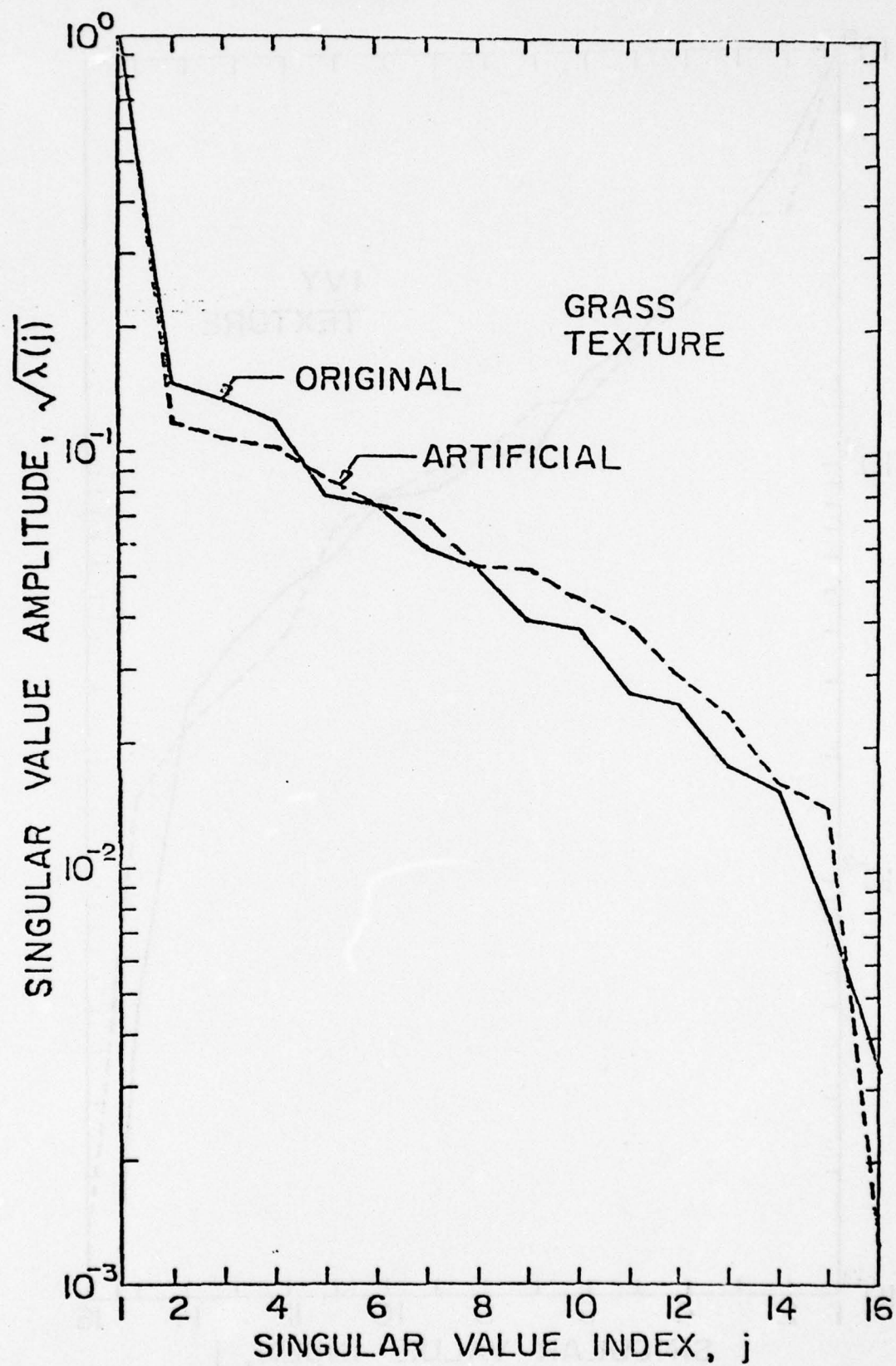


Figure 4. Singular Values of Natural and Artificial Grass Texture.

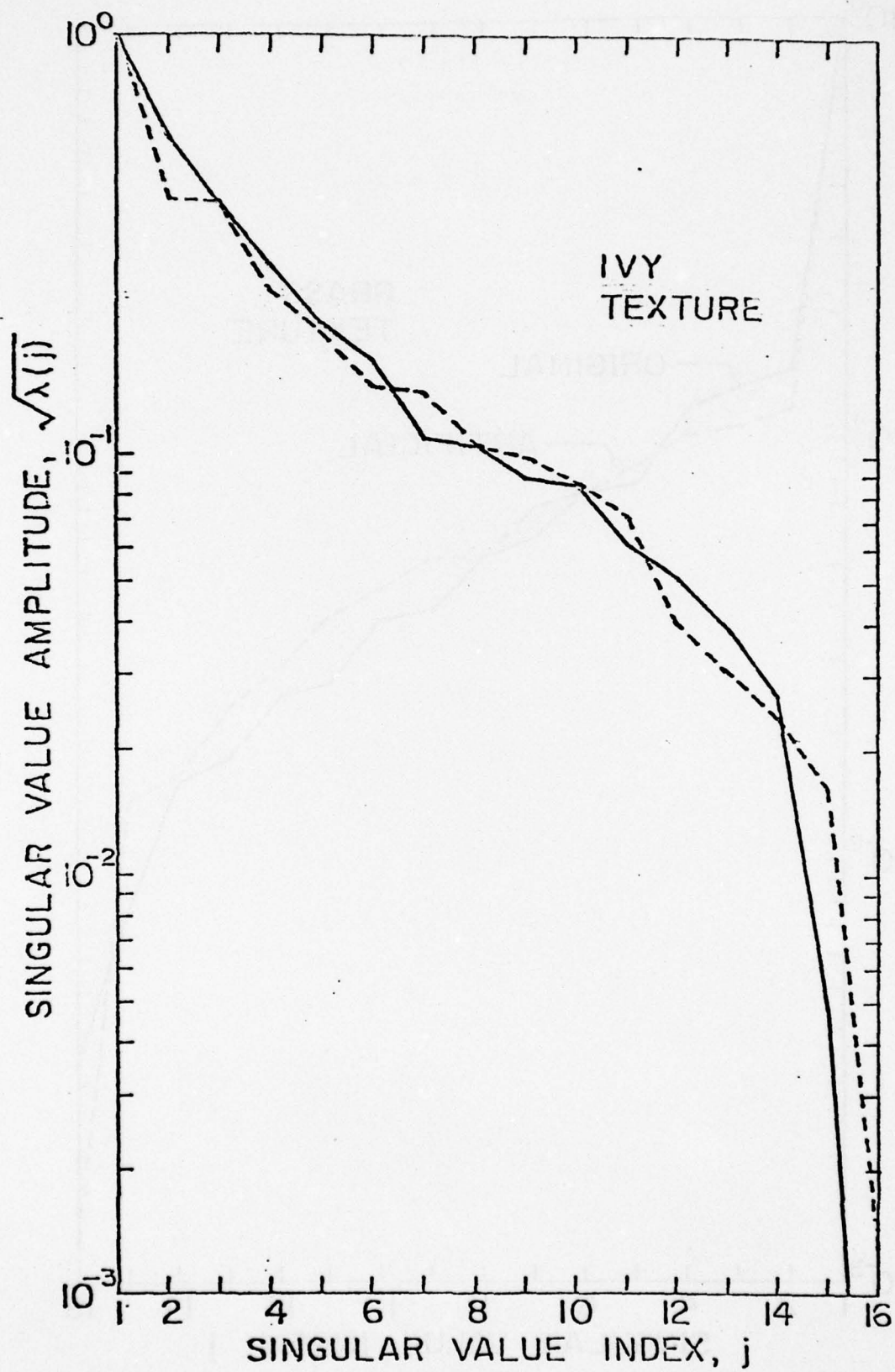


Figure 5. Singular Values of Natural and Artificial Ivy Texture.

grass and ivy are seen to be significantly different, but the distributions of the natural and artificial pairs are quite close.

References

1. G.H. Golub and C.L. Patterson, "Singular Value Decomposition and Least Squares Solutions," Numerical Mathematics, Vol. 14, 1970, pp. 403-420.
2. H.C. Andrews and C.L. Patterson, "Outer Product Expansions and Their Uses in Digital Image Processing," American Mathematical Monthly, Vol. 1, No. 82, January 1975, pp. 1-13.
3. H.C. Andrews and C.L. Patterson, "Outer Product Expansions and Their Uses in Digital Image Processing," IEEE Transactions on Computers, Vol. C-25, No. 2, February 1976, pp. 140-148.
4. W.K. Pratt, "Vector Space Formulation of Two-Dimensional Signal Processing Operations," Journal Computer Graphics and Image Processing, Vol. 4, No. 1, March 1975, pp. 1-24 (Academic Press, New York).
5. S.S. Wilks, Mathematical Statistics, Wiley, New York, 1962.
6. J. Wishart, "The Generalized Product Moment Distribution in Samples from a Normal Multivariate Population," Biometrika, Vol. 20A, 1928, pp. 32-52.
7. A.T. James, "The Distribution of the Latent Roots of the Covariance Matrix," Annals Mathematical Statistics, Vol. 31, 1960, pp. 151-158.

3. Image Processing Projects

The image processing projects described in the succeeding pages represent a broad variety of research topics in which more fundamental signal processing principles are utilized. The first contribution presents a new way of generating binary phase holograms via the digital computer. Such holograms are then recorded on hardcopy devices (see Section 5) and after film development are played back in the Institute optical laboratories. The second contribution in this section presents rather encouraging results on the blind deconvolution process of a posteriori restoration. This project is currently coming to fruition with quite improved results for arbitrary space invariant distortions. A radar imaging task is next presented in which multi-frequency radar returns are coherently processed to form images of aircraft mounted on a turntable. Previous reports included the analysis associated with this theory and this report presents the images resulting from such processing. This section closes with a brief summary of perceptual model color image coding. Results are presented in which 24:1, 48:1, and 96:1 compressions are obtained with good visual color quality maintained.

3.1 Double Phase Holograms, A New Way of Generating Binary Holograms (Supported by NSF Grant ENG-76-15318 and AFOSR under Contract AFOSR-77-3285)

Chung-Kai Hsueh and Alexander A. Sawchuk

Of all the classes of computer generated holograms, binary holograms are the easiest to make. Having the Fourier transform of the object, only a plotter or a binary

display device and photo-reduction equipment are required to make a binary hologram. They also have the advantage of superior signal-to-noise ratio due to the binary transmission nature of the hologram. The efficiency of a binary hologram can also be increased by bleaching.

The major difficulty with digital holography is in finding a method to record complex transmission values on a photographic or other optical recording medium which can store non-negative real values. Area modulation and grey levels are used to represent amplitude, and various methods of encoding the phase by position modulation [1] within a Fourier transform resolution cell, or by the use of subcells [2-3] have been developed.

Since any complex value can be expressed by the summation of non-negative multiples of +1, -1, +j and -j Lee [2] divided each cell into four subcells with grey levels in each subcell to represent the amplitude of each vector component. Burckhardt [3] further found that instead of using four vectors to represent a complex value, three vectors pointing in the directions 0, $2\pi/3$ and $-2\pi/3$ can be used to represent any complex value.

The Double Phase Hologram (DPH) introduced here is based on the principle that any complex transmission value $Ae^{j\theta}$ inside the unit circle ($0 \leq A \leq 1$) can be decomposed into the sum of two constant magnitude vectors as expressed by [4]

$$H = Ae^{j\theta} = \frac{1}{2}e^{j(\theta+\psi)} + \frac{1}{2}e^{j(\theta-\psi)} \quad (1)$$

where

$$\psi = \cos^{-1}A, \quad 0 \leq \psi \leq \pi/2 \quad (2)$$

In this expression we have dropped the (u,v) coordinates of the transform domain for simplicity. The corresponding coordinates in the image plane will be designated (x,y) . The two terms on the right side of Eq. (1) are designated H and H^* respectively, and these are shown in Fig. 1. The decomposition can be applied to represent a complex value in a binary hologram by using the two phase quantities instead of the three non-negative quantities in a Burckhardt hologram.

These two phase quantities are implemented by detour phase as shown in Fig. 2. Each Fourier component on the hologram consists of two subcells with each subcell having an open slit of fixed area in the opaque background. The position is determined by the decomposed phases from

$$d_1 = d(\theta + \psi)/2\pi \quad (3)$$

and

$$d_2 = d(\theta - \psi)/2\pi \quad (4)$$

where d is the size of one Fourier resolution cell. If we consider the subcells as a whole, then each component has the correct complex value in the direction of the first diffraction order along the v direction. Therefore the image appears at the first diffraction order in the back focal plane of the lens. This type of hologram is called a Double Phase Hologram (DPH).

From Eq. (2) the angle ψ is related to the amplitude A by

$$A = \cos\psi \quad 0 \leq \psi \leq \pi/2 \quad (5)$$

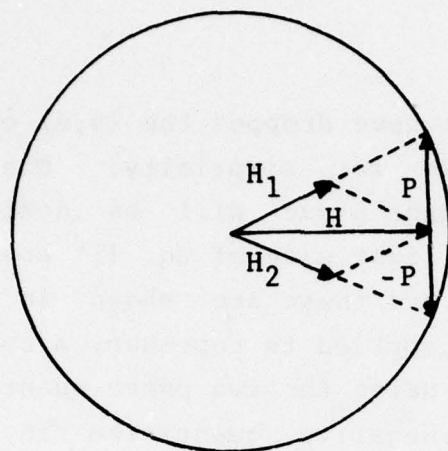


Figure 1. Decomposition of a vector in the unit circle into two constant magnitude vectors H_1 and H_2 .

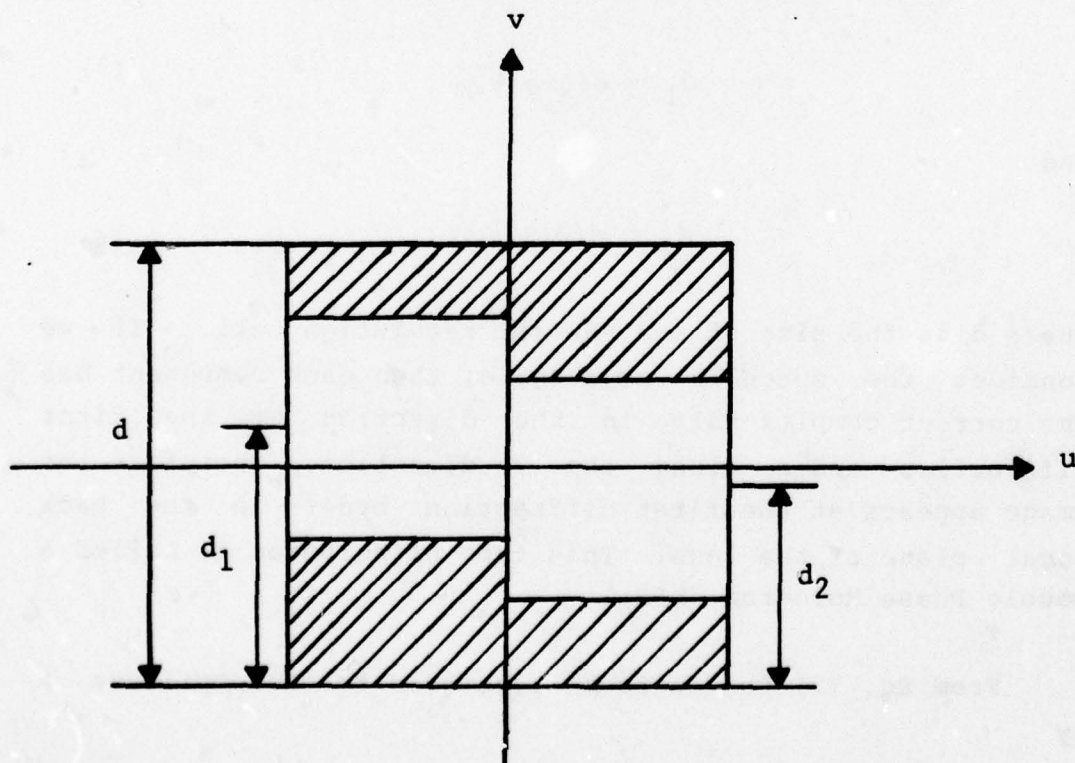


Figure 2. Implementation of a Double Phase Hologram.

and its derivative is given by

$$\frac{dA}{d\psi} = \sin\psi \quad 0 \leq \psi \leq \pi/2 \quad (6)$$

For the DPH ψ is quantized to a certain level $\hat{\psi}$ and the amplitude becomes

$$\hat{A} = \cos\hat{\psi} \quad (7)$$

rather than A . As shown in Eq. (6), the derivative of A is an increasing function in the defined range. Therefore the same amount of quantization error in $\hat{\psi}$ results in smaller error in \hat{A} when A is close to 1 (ψ close to 0) and larger error in \hat{A} when A is close to 0 (ψ close to $\pi/2$). This indicates that it is better to have an uniform spectrum in order to reduce the quantization error. An object with zero phase usually has a very large dynamic range and energy is concentrated at low frequency components. To spread out the spectrum we can either impose a random phase on the object or use deterministic diffusers [5]. Other iterative phase coding techniques [6-7] can also be used to smooth out the spectrum in a desired form and therefore reduce the quantization error in the DPH.

It has been found that although the phase coding techniques reduce the dynamic range of the Fourier transform, they also introduce speckles due to aliasing error. The space-bandwidth iterative method with constrained bandwidth [8] has been used to reduce this noise. Here we can use a similar idea to implement a DPH. First of all a Fourier transform with constrained bandwidth is obtained by the iterative method. Then only the center part of the transform is used in plotting a hologram. By doing this we can reduce the speckle noise while preserving

the same space-bandwidth product. The uniform transform required for correct phase quantization in a DPH is also maintained.

The implementation of a DPH is based on the decomposition of a vector into two other constant vectors as given by Eqs. (1) and (2). By defining

$$H + P = 2H_1 \quad (8)$$

and

$$H - P = 2H_2 \quad (9)$$

where P is a vector called the parity term shown in Fig. 1, the decomposition of Eqs. (1) and (2) can be rewritten as [9]

$$H = H_1 + H_2 = \frac{1}{2}(H+P) + \frac{1}{2}(H-P) \quad (10)$$

If H_1 and H_2 were overlapped at the same position, P and $-P$ will cancel out and the correct transform is obtained. However due to the method of implementing the DPH, these two terms do not exactly cancel out, generally resulting in parity noise.

Let us denote $h(x,y)$ and $p(x,y)$ as the inverse Fourier transforms of $H(u,v)$ and $P(u,v)$ respectively. Using the configuration in Fig. 2, it can be shown that the reconstruction at the first diffraction order along the y direction is given by

$$O(x,y) = h_s(x,y) + p_s(x,y) \quad (11)$$

with

$$h_s(x,y) = d^2 W_h(x,y) [h(x,y) * \text{comb}(dx, dy)] \quad (12)$$

$$p_s(x,y) = j d^2 W_p(x,y) [p(x,y) * \text{comb}(dx, dy)] \quad (13)$$

where

$$W_h(x,y) = \text{sinc}(dx, dy) \quad (14)$$

and

$$W_p(x,y) = -\text{sinc}(dx/2, dy) \sin(\pi dx/2) \quad (15)$$

are the weighting functions for the desired function and the noise term.

Both $W_h(x,y)$ and $W_p(x,y)$ are drawn in Figs. 3 and 4 respectively. At $x = 0$ the noise term $p(x,y)$ is strongly attenuated by $\sin(\pi dx/2)$. Since $h(x,y)$ is space-limited we may assume that $p(x,y)$ is also space-limited to a small region. Therefore in the region of interest (zero order) noise is less noticeable. On the other hand at $x = 1/d$ (first order) $\sin(\pi dx/2)$ reaches its maximum and $\text{sinc}(dx/2)$ does not fall off too much while $h(x,y)$ is strongly attenuated by $\text{sinc}(dx)$ at $x = 1/d$. Therefore the error term dominates at the first diffraction order.

Each cell can be further divided into more subcells as long as they sum up with the correct transform value. Suppose there are $2n$ subcells in each cell with H_1 and H_2 plotted alternatively. We can proceed as before and obtain Eqs. (11)-(14) with

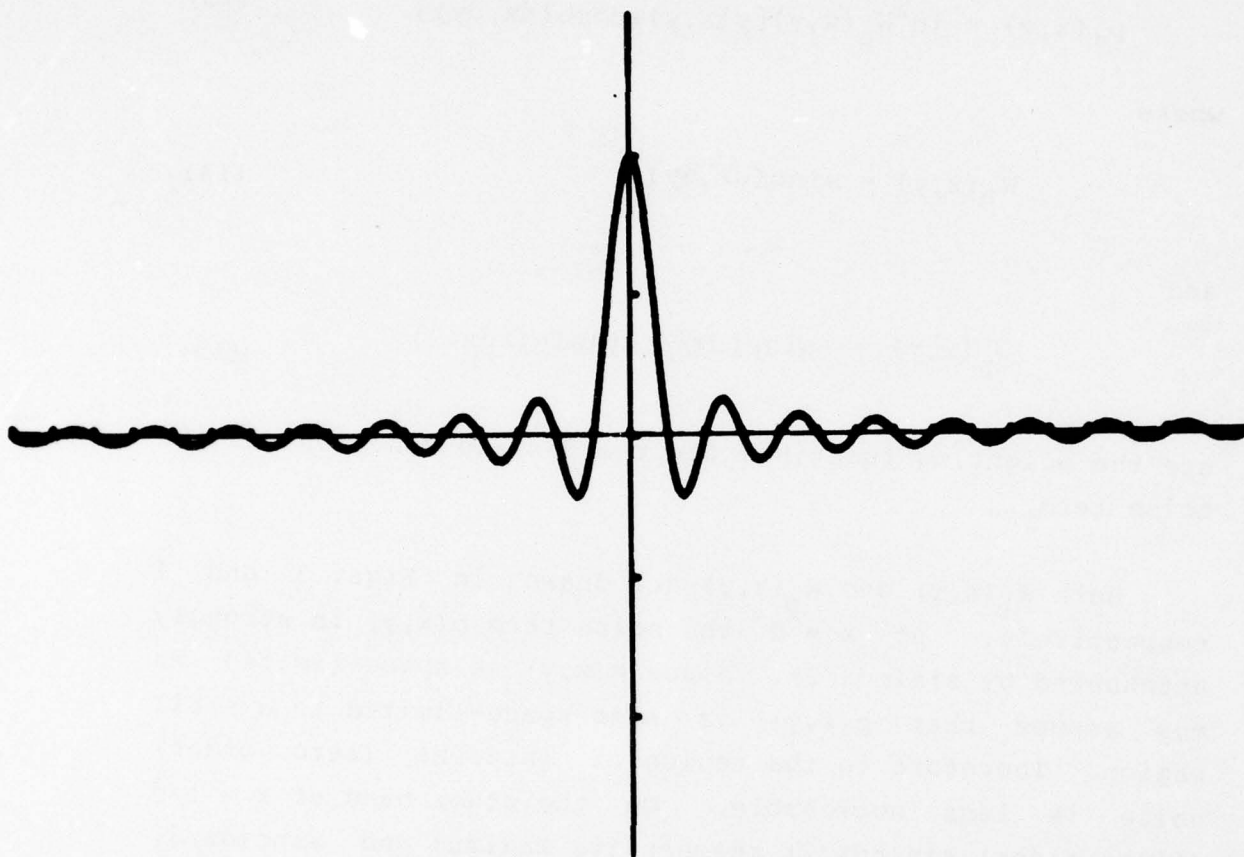


Figure 3. Weighting function W_h for the reconstruction.

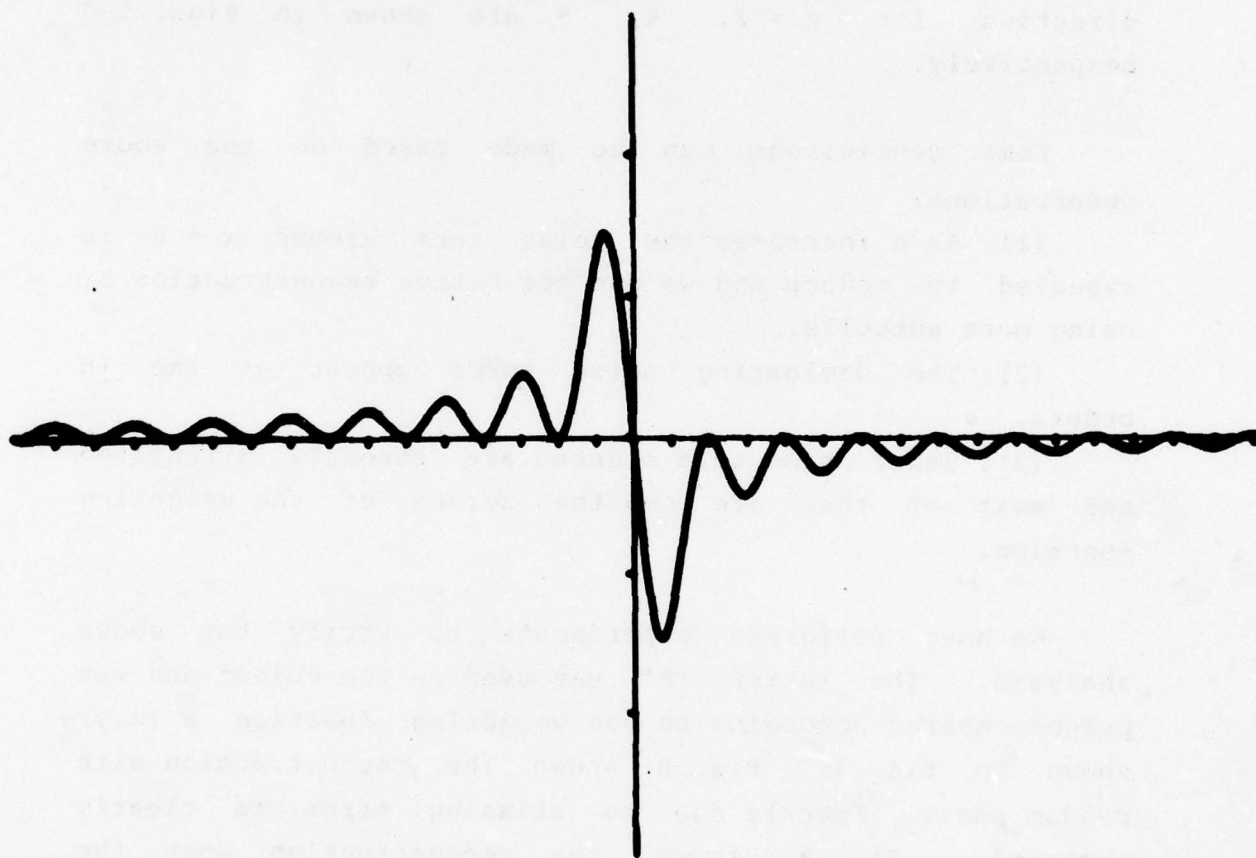


Figure 4. Weighting function W_p for the noise term (2 subcells).

$$W_p(x,y) = \frac{(-1)^n}{n} \text{sinc}(dx/2n, dy) \{ \sin\phi - \sin 3\phi + \dots + (-1)^{n-1} \sin(2n-1)\phi \} \quad (16)$$

where $\phi = \pi dx/2n$. The curves of $W_p(x,y)$ along the x direction for $n = 2, 4, 8$ are shown in Figs. 5-7 respectively.

Some conclusions can be made based on the above observations:

(1). As n increases the noise term around $x = 0$ is expected to reduce and we can get better reconstruction by using more subcells.

(2). The dominating noise terms appear at the $\pm n$ orders.

(3). Other noise term orders are strongly attenuated and most of them are on the zeroes of the weighting function.

We have performed experiments to verify the above analysis. The letter 'P' was used as the object and was precompensated according to the weighting function $W(x,y)$ shown in Fig. 3. Fig. 8 shows the reconstruction with random phase. Speckle due to aliasing error is clearly observed. Fig. 9 shows the reconstruction when the space-transform iterative method with constrained bandwidth is used and only the central part is plotted. The image quality is considerably improved.

References

1. A.W. Lohmann and D.P. Paris, "Binary Fraunhofer Holograms, Generated by Computers," Applied Optics, 6, pp. 1739-1748 (1967).

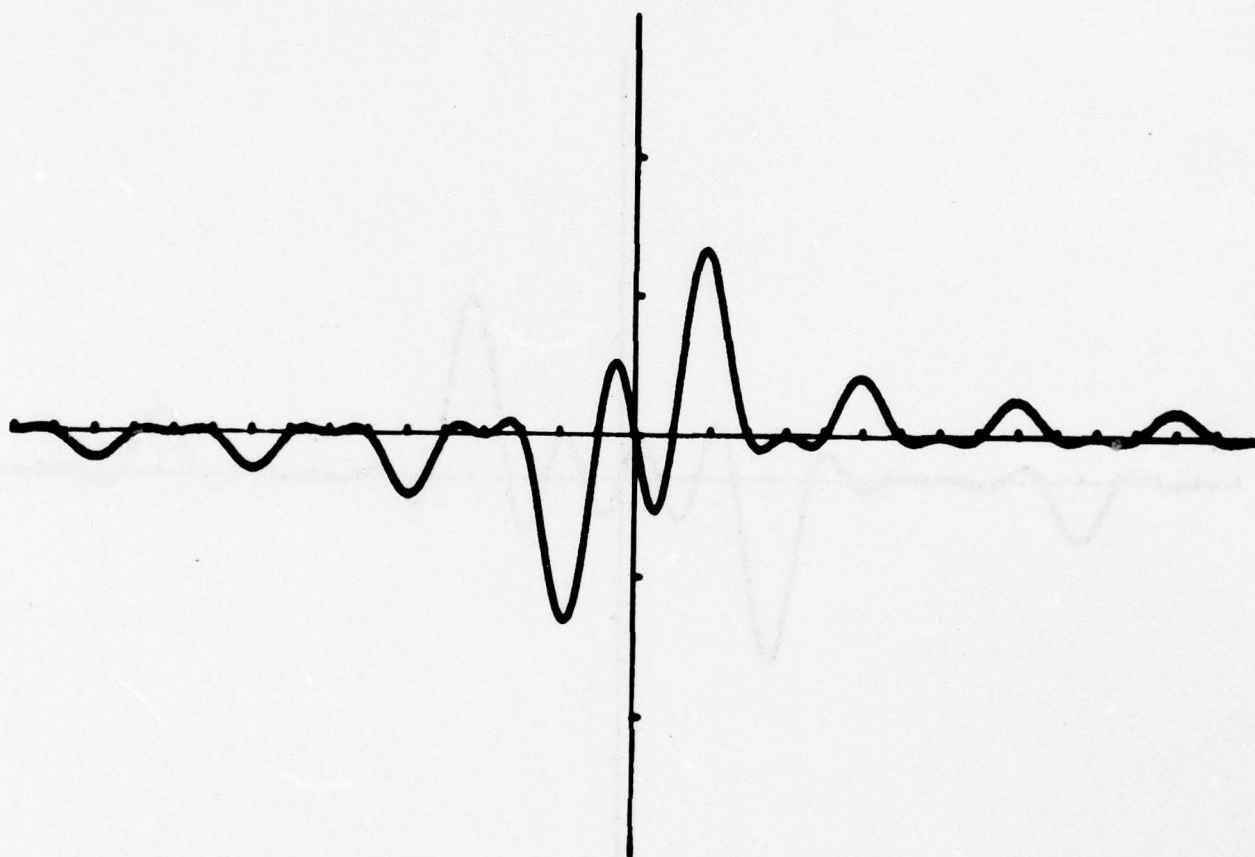


Figure 5. Weighting function W_p for the noise term (4 subcells).

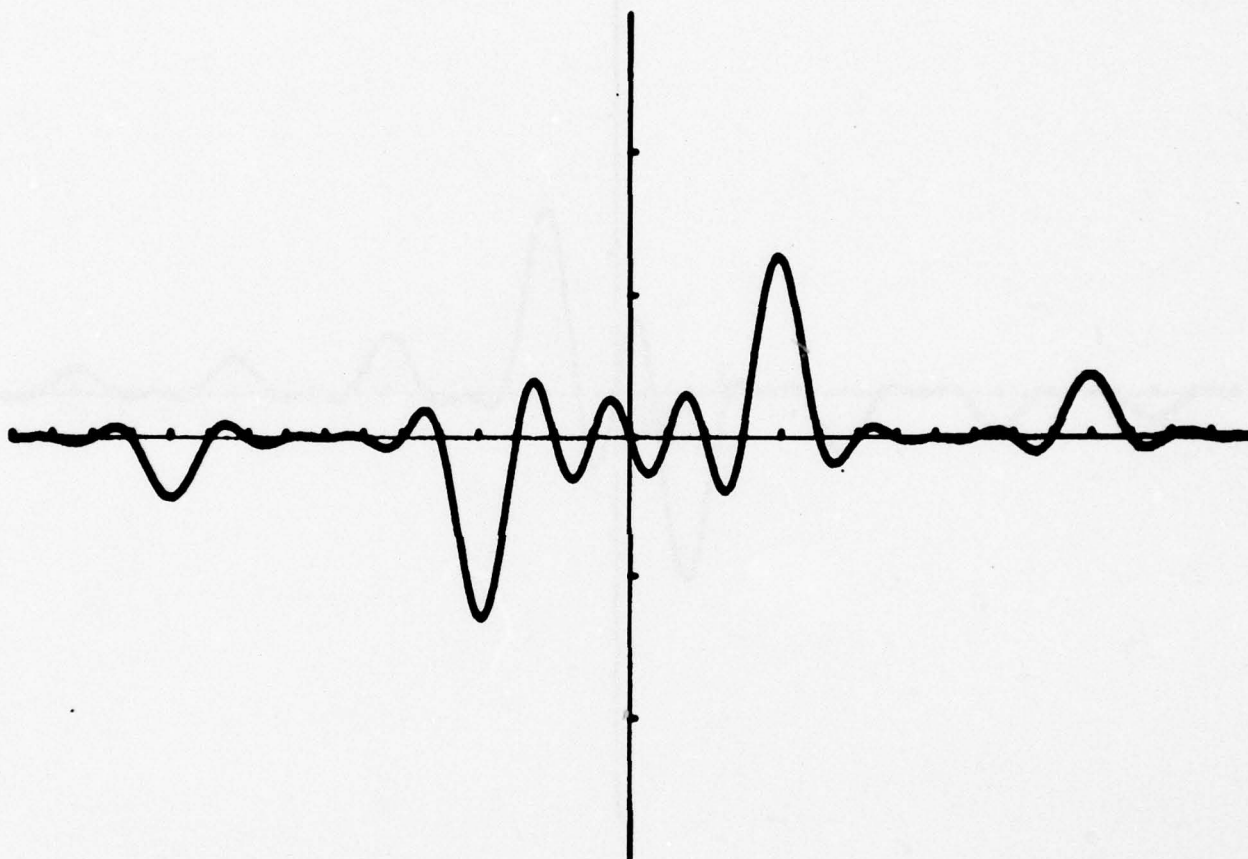


Figure 6. Weighting function W_p for the noise term (8 subcells).

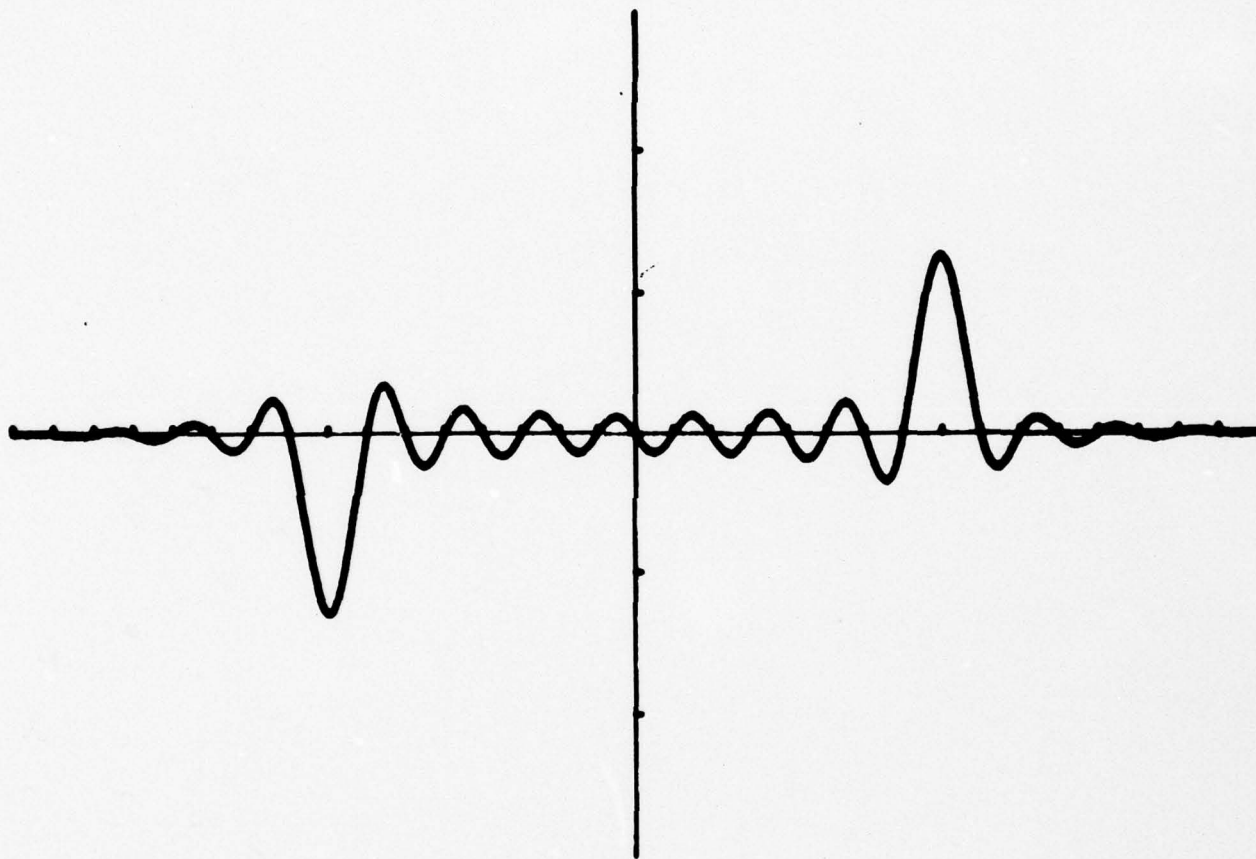


Figure 7. Weighting function W_p for the noise term (16 subcells).

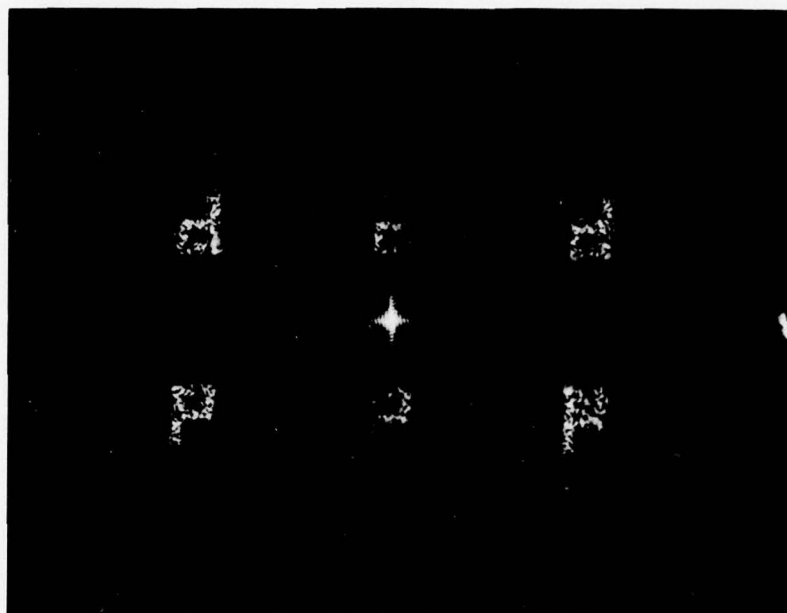


Figure 8. Reconstruction of letter P with random phase.

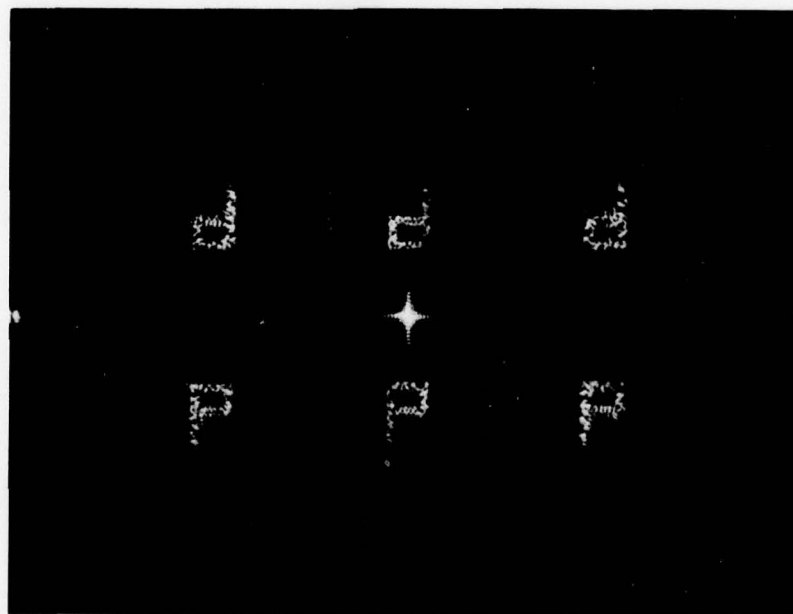


Figure 9. Reconstruction with constrained bandwidth.

2. W.H. Lee, "Sampled Fourier Transform Hologram Generated by Computer," Applied Optics, 9, pp. 639-643 (1970).
3. C.B. Burckhardt, "A Simplification of Lee's Method of Generating Holograms by Computer," Applied Optics, 9, p. 1949 (1970).
4. A.A. Sawchuk and C.K. Hsueh, "On-Axis Optical Filtering System," USCIPI Report 770, pp. 151-161 (Sept. 1977).
5. W.J. Dallas, "Deterministic Diffusers for Holography," Applied Optics, 12, pp. 1179-1187 (1973).
6. N.C. Gallagher and B. Liu, "Method for Computing Kinoforms that Reduces Image Reconstruction Error," Applied Optics, 12, pp. 2328-2335 (1973).
7. J.R. Fienup, "Improved Synthesis and Computational Methods for Computer-Generated Holograms," Ph.D. Thesis, Department of Electrical Engineering, Stanford University (1975).
8. J.P. Allebach, N.C. Gallagher and B. Liu, "Aliasing Error in Digital Holography," Applied Optics, 15, pp. 2183-2188 (1976).
9. M. Severcan, "Computer Generation of Coherent Optical Filters with High Light Efficiency and Large Dynamic Range," Ph.D. Thesis, Stanford University (1973).

3.2 A Technique of A Posteriori Restoration--Results of a Computer Simulation

John Morton

Illustrated in Figure 1 is the blurring point-spread-function (PSF) used in the computer simulation reported herein.

Figure 2 compares the \log_{10} of the magnitude of the optical transfer function (OTF) with the corresponding estimate. For comparison purposes values of the log of the magnitude below 2 have been set to 2. The estimate used the method of Cannon [1].

Figure 3 compares the phase of the OTF with two estimates. Both estimates used the method as previously reported [2-4]; the differences in the two estimates is that Estimate 1 relaxed one assumption implicit in the method, whereby Estimate 2 did not. The comparisons for the phase of the OTF on the frequency axes are given in Figure 4.

Figure 5 presents the image degraded by the PSF in Figure 1 together with restorations assuming different a priori knowledge. The assumptions underlying the restorations are given in Table 1.

Upon comparison of the magnitude of the optical transfer function and the estimate of the magnitude, it is clear that the estimate adequately retains the essential features.

Upon comparison of the phase of the optical transfer function and the estimates of the phase, it is evident that phase Estimate 1 is reasonable near the axes and not as accurate at a distance from the axes. Because images in general have little energy concentration in those frequency

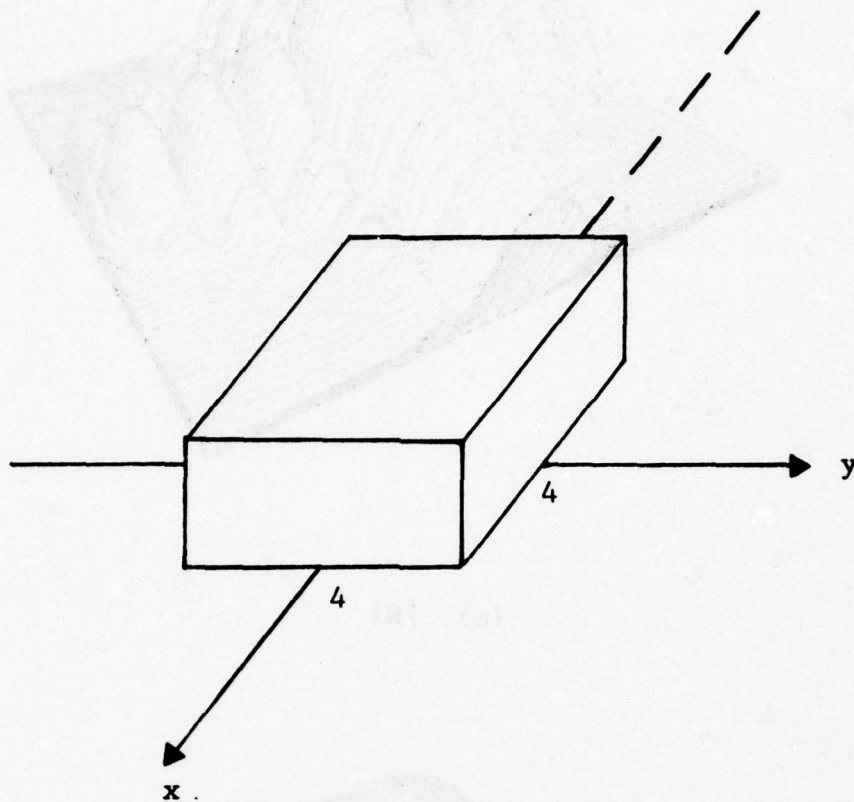
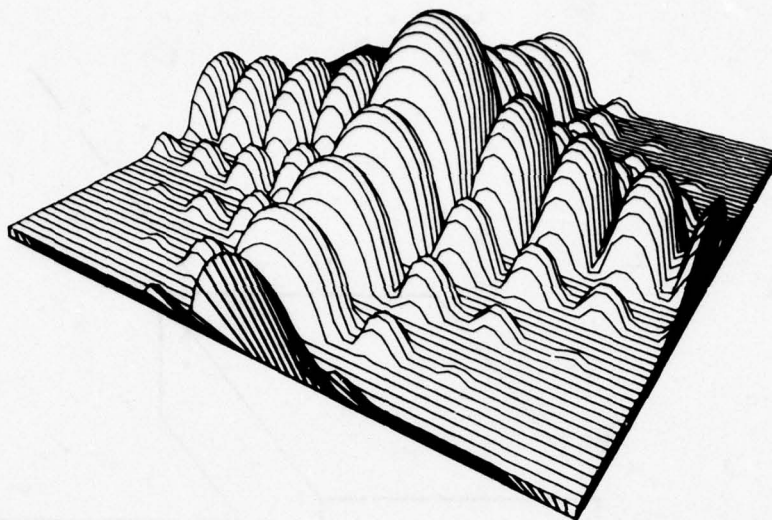
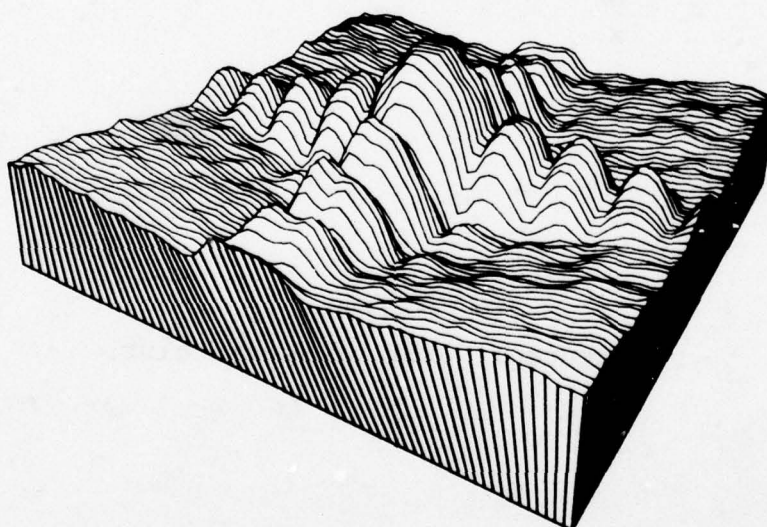


Figure 1. PSF of square blur.

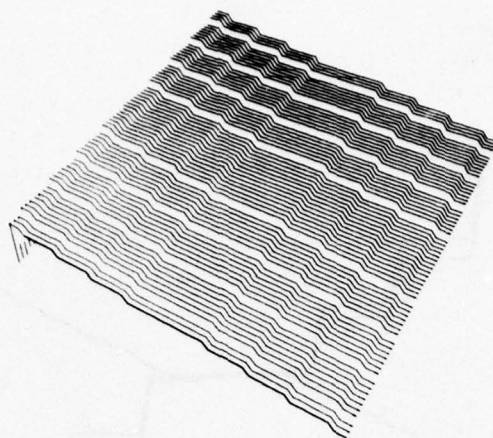


(a) $|H|$

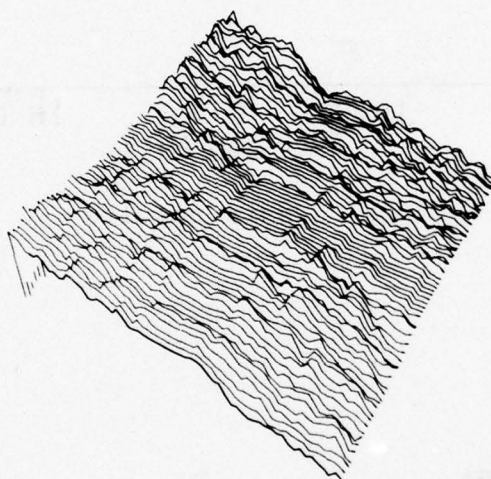


(b) $|\hat{H}|$

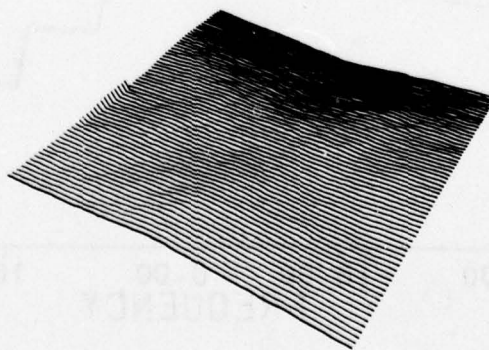
Figure 2. Comparison of magnitude of OTF and estimate.



(a) θ

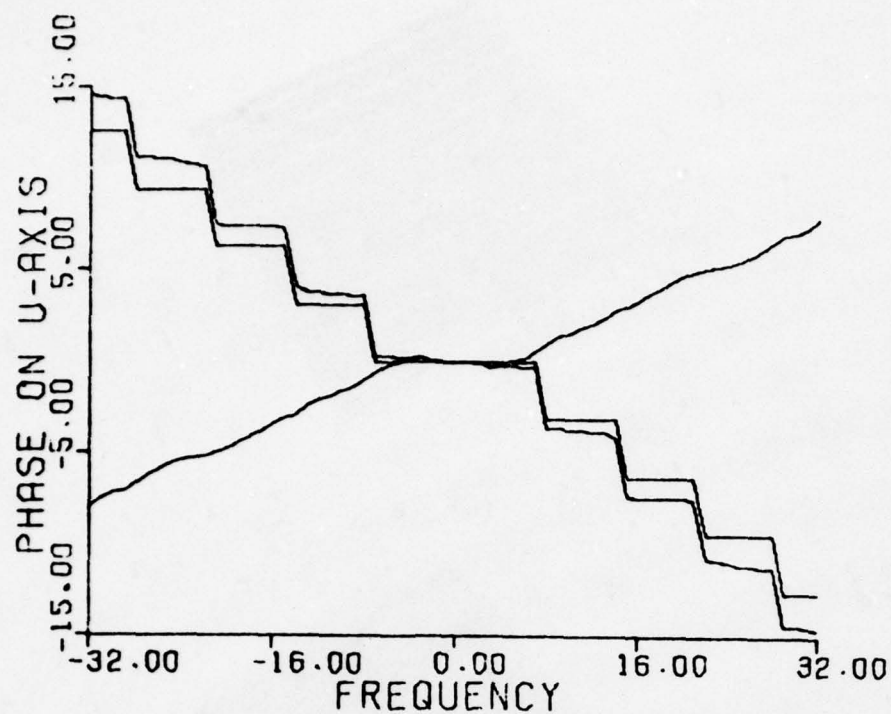


(b) $\hat{\theta} = \text{Estimate 1}$

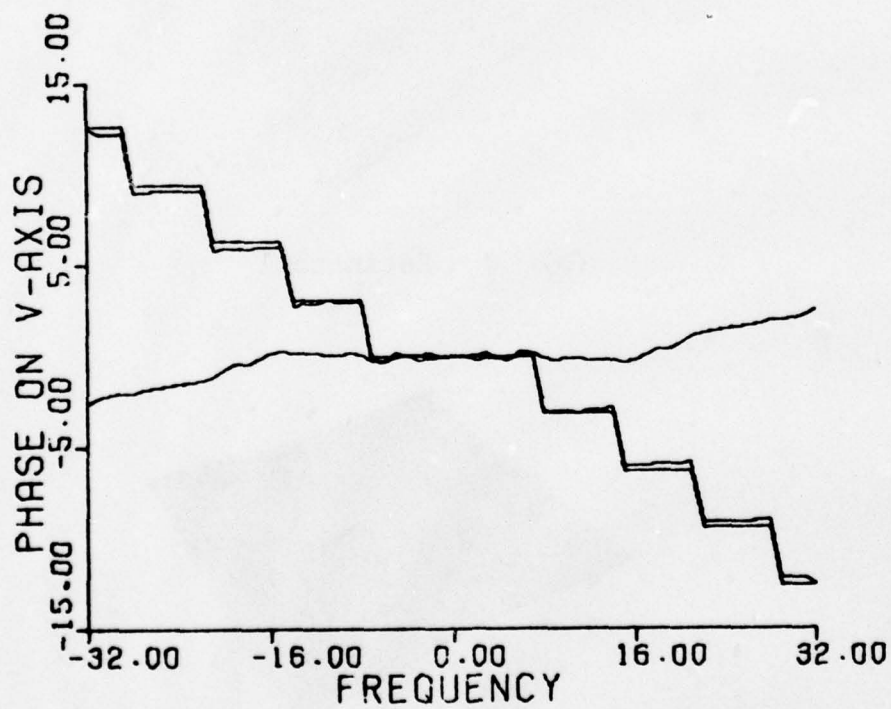


(c) $\hat{\theta} = \text{Estimate 2}$

Figure 3. Comparison of phase of OTF and estimates.

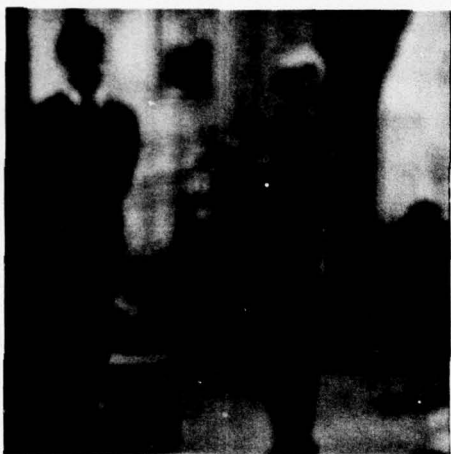


(a) u axis results



(b) v axis results

Figure 4. Comparison of phase of OTF and estimates.



(a) degraded image



(b) magnitude of OTF given
phase of OTF given



(c) magnitude of OTF given
phase of OTF estimated as 0



(d) magnitude of OTF given
phase estimate = Estimate 2

Figure 5. Restorations.



(e) magnitude of OTF estimated
phase of OTF given



(f) magnitude of OTF estimated
phase of OTF estimated as 0



(g) magnitude of OTF estimated
phase estimate = Estimate 2



(h) magnitude of OTF estimated
phase estimate = Estimate 2

Figure 5 (continued)

<u>Figure</u>	<u>OTF magnitude</u>	<u>OTF phase</u>	<u>Power Spectrum of Undegraded image</u>
b	given	given	given
c	given	estimated as 0	given
d	given	Estimate 2	given
e	estimated	given	estimated
f	estimated	estimated as 0	estimated
g	estimated	Estimate 2	estimated
h	estimated	Estimate 1	estimated

Table 1. Key to restorations.

components away from the axes, the poorer results away from the axes are not very important from a restoration point of view.

From the phase comparisons it may be noted that phase Estimate 2 is not particularly good.

Comparing restorations, it is apparent that the estimate of the magnitude of the OTF is reasonable to the extent of achieving an adequate restoration. In fact, the visual quality of the restoration using the estimate of the magnitude of the OTF seems to be superior to the restoration given knowledge of the magnitude of the OTF. In general, upon restoration, the frequencies in the vicinity of the zeros of the OTF will be magnified the most by the Wiener filter. If one knows the magnitude of the OTF exactly, at times the restoration will contain periodic artifacts corresponding to those frequencies that were boosted the most by the filter. On the other hand, the inaccuracies inherent in the estimate of the magnitude of the OTF cause the frequencies in the vicinity of the zeros to be boosted less than the boost assuming knowledge of the magnitude of the OTF. As a result, when assuming the estimate of the magnitude of the OTF, the periodic artifacts are missing.

Although the visual quality of the restoration is usually superior when using the estimate of the magnitude of the OTF, in terms of minimum mean squared error, restoration b is superior.

Upon comparing restorations c and d and restorations f and g, it is generally agreed that Estimate 2 of the phase is not significantly better with respect to restoring than an estimate of zero for the phase.

The superiority of Estimate 1 of the phase is evident. For example, note less distortion in the face of the man in h opposed to f and g. Also note the streaks of light in the pants of the man and across the chest and below the knees of the woman in f and g. These anomalies are absent in h.

References

1. T.M. Cannon, "Digital Image Deblurring by Nonlinear Homomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSC-74-091, August 1974.
2. Semiannual Technical Report, Harry C. Andrews-project director, September 30, 1976, USCIPI Report 720.
3. Semiannual Technical Report, Harry C. Andrews-project director, March 31, 1977, USCIPI Report 740.
4. Semiannual Technical Report, Harry C. Andrews-project director, September 30, 1977, USCIPI Report 770.

3.3 Turntable Radar Imaging

Chung Ching Chen and Harry C. Andrews

Introduction

Radar systems use active device to radiate radio frequency signals and receive their returns. Various information about the targets can be inferred by processing those echoed data [1].

In two dimensional radar imaging system the two geometric coordinates are usually range and azimuth [2]. Range is the direction along which the signal is transmitted and received. Azimuth is the direction orthogonal to the range direction and along the relative motion of radar and targets. The range information is obtained by timing the radar return by either short pulse or pulse compression techniques. In the classical radar systems azimuth information is provided by using very long antenna whose illuminating pattern is very narrow along azimuth, because

$$\beta = \lambda/L \quad (1)$$

where β is the effective angle of antenna pattern, λ is the wavelength and L is the length of the antenna. To achieve small β comparable to the range resolution usually requires impractically long antennas. Coherent radar systems get around this problem by synthesizing without actually implementing long antennas created by the motion of the targets and the radar [3], [4], [5].

In this paper we study a radar system which has mathematical similarity with the computer aided tomography (CAT) used in medical applications. Degrees of freedom (DOF) [6] of such a system is derived for the purpose of efficient computation and minimal data storage requirement. Physical differences between the two systems are analyzed and incorporated in the tailored reconstruction algorithm unique to the underlying radar system to reconstruct the target images.

Experiments are done in support of the theory developed in this work.

The RAT SCAT Facility - Data Acquisition

The RAT SCAT (RADar Target SCATter Site) [7] facility is used to obtain the data for our processing. A target to be imaged is placed on a rotator at a distance r_0 from the radar to the rotation center as in Fig. 1. The reference sphere S sits at distances r_1 from the radar R and r_2 from the rotation center C. As shown in Fig. 2 there are two rectangular coordinates systems. (ξ, η) and (x, y) , with the later system making an angle θ from the former one. We assume that (ξ, η) is fixed with the target and (x, y) is fixed with the radar or ground. The radar radiates the same set of step frequencies $\{f_k\}$ at discrete aspect angles from the set $\{\theta_i\}$. Normally the step frequency Δf and step angle $\Delta\theta$ are fixed, and we assume so in our case. The local oscillator at S takes the signal from R to S and beats it with the signal reflected from the target to form in-phase and quadratic-phase data components.

Target Reflectivity Function

Define the reflectivity function $f(\xi, \eta)$ of the target as the ratio of the received signal due to the target point at (ξ, η) and the transmitted signal, as in Fig. 2. The metallic target looks specular to the radar because the wavelength λ used is very small compared to the target size. There is also shadowing effect because of the non convexity of the target surface. Thus $f(\xi, \eta)$ is actually a function of θ . However we'll assume for the moment that $f(\xi, \eta)$ is independent of θ and we'll derive the PSF as well as DOF of such a system which offer much insight into the mathematical properties of the systems.

PSF Of The Imaging System

In Fig. 1 let's assume that angles α and β are virtually zero. The distance between the target point (ξ, η)

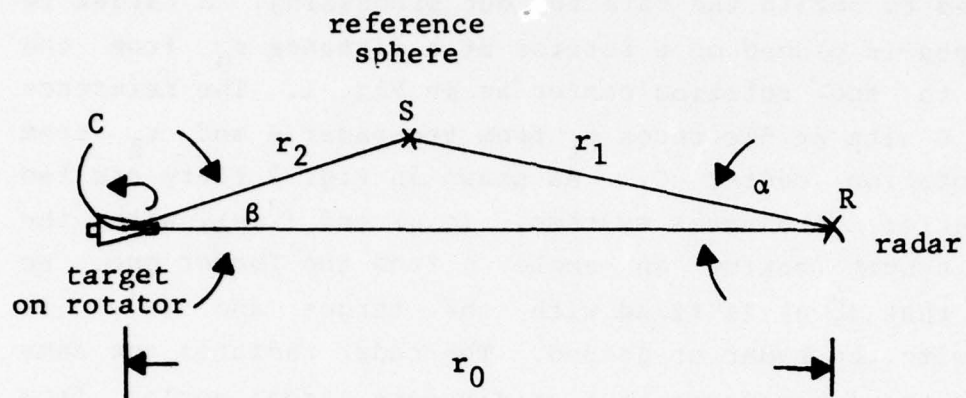


Fig. 1. relation among radar, target and reference sphere

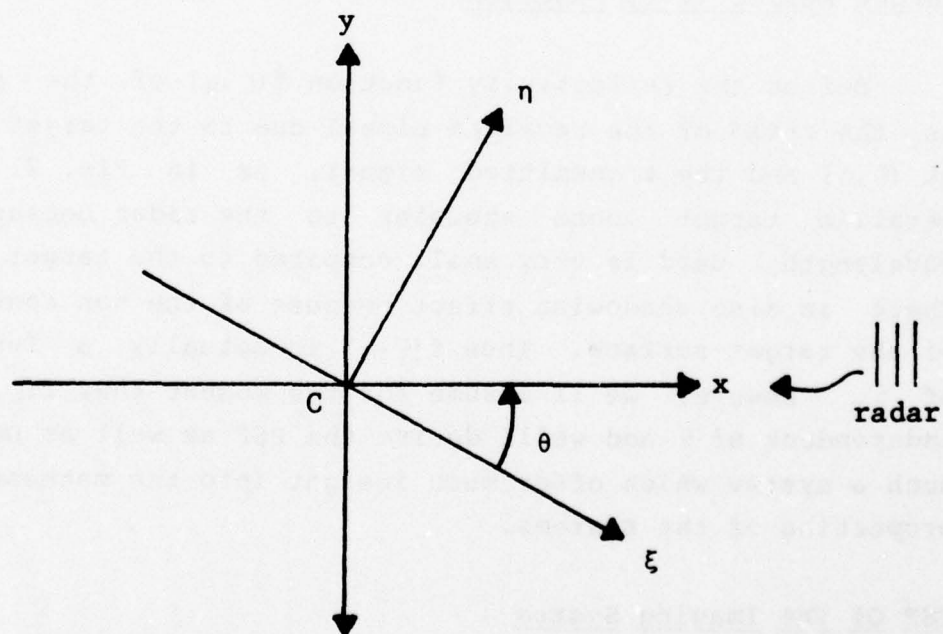


Fig. 2. geometry of two coordinate systems
 (ξ, η) : target (x, y) : ground or radar

and the radar is

$$\begin{aligned} r(r_0, \theta, \xi, \eta) &= [(r_0 - x)^2 + y^2]^{\frac{1}{2}} \\ &\approx (r_0 - x), \end{aligned} \quad (2)$$

where we have assumed $r_0 - x \gg y$ for all target points (x, y) and

$$\begin{cases} x = \xi \cos \theta + \eta \sin \theta \\ y = \eta \cos \theta - \xi \sin \theta \end{cases} \quad (3)$$

Differentiating (1) with respect to θ and incorporating Eq. (2) one gets

$$\frac{\partial r(r_0, \theta, \xi, \eta)}{\partial \theta} \approx \xi \sin \theta - \eta \cos \theta = -y \quad (4)$$

where $r_0 - x \approx r(r_0, \theta, \xi, \eta)$ has been assumed. It is observed from Eqs. (2) and (4) that the lines of constant range and the lines of constant doppler are parallel to the y, x , axis, respectively.

Define

$$g(x, \theta) = \int_{-a}^a f(\xi, \eta) dy \quad (5)$$

where a is the maximal radial extent of the target, and

$$f_k(t) = A \cos(2\pi f_k t + \phi) \quad k = 1, 2, \dots, M \quad (6)$$

which are the signals transmitted from the radar. The signals returned to S from the target line mass at range $r_0 - x$ is

$$z(x, i, k, t) = B g(x, \theta_i) \cos \left[2\pi f_k \left(t - \frac{r_0 + r_2 - 2x}{c} \right) + \phi \right] \quad (7)$$

where the signal power, propagation decay and the reflectivity phase have been absorbed into the complex constant B. The total return is

$$\begin{aligned} z(i, k, t) &= \int_{-a}^a z(x, i, k, t) dx \\ &= B \int_{-a}^a g(x, \theta_i) \cos \left[2\pi f_k \left(t - \frac{r_0 + r_2 - 2x}{c} \right) + \phi \right] dx \quad (8) \end{aligned}$$

The signal along path RS received by S assumes the form

$$r(k, t) = C \cos \left[2\pi f_k \left(t - \frac{r_1}{c} \right) + \phi \right] \quad (9)$$

Beating signals (7) and (8) and appropriately filtering the result in in-phase and quadratic-phase components

$$\begin{aligned} I(i, k) &= D \int_{-a}^a g(x, \theta_i) \cos \left\{ 2\pi f_k \frac{2x}{c} + 2\pi f_k \frac{r_1 - r_0 - r_2}{c} \right\} dx \\ Q(i, k) &= -D \int_{-a}^a g(x, \theta_i) \sin \left\{ 2\pi f_k \frac{2x}{c} + 2\pi f_k \frac{r_1 - r_0 - r_2}{c} \right\} dx \quad (10) \end{aligned}$$

were not for the linear phase factor Eq.(10) would be the Fourier components of the shadow gram $g(x, \theta_i)$ at angle θ_i and frequency $\frac{2f_k}{c}$, [8], [9], [11] and would take values on the ring area as in Fig. 3 just like that in the tomography. DOF of the tomographic system has been analyzed in detail by McCaughey and Andrews [10] with the same geometry as shown in Fig. 4. Because of the mathematical similarity we'll apply some of the results in [10] to find the DOF of the raw radar data, keeping in mind the physical differences such as the angle-dependence of reflectivity function mentioned

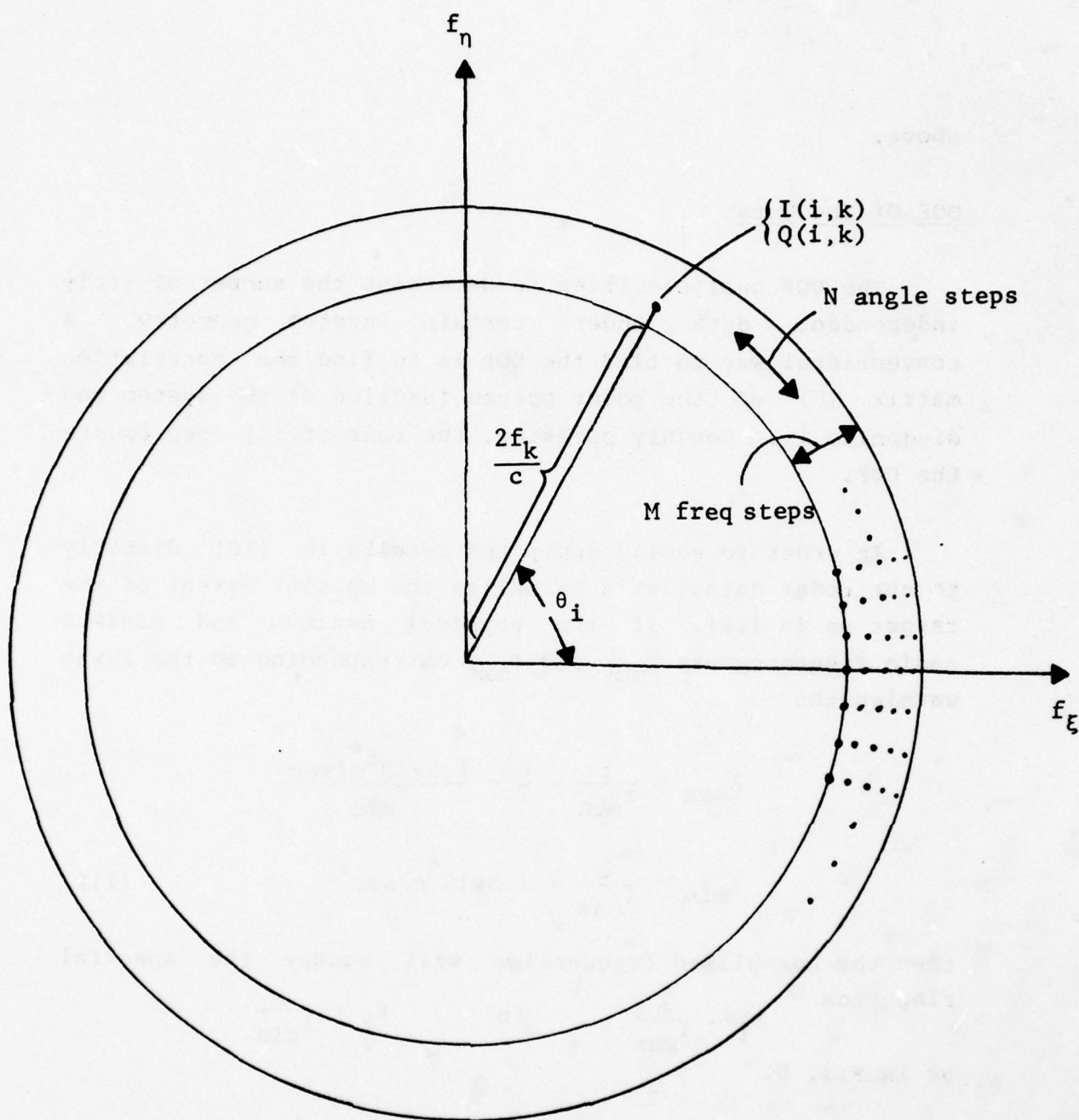


Fig. 3. data in frequency domain

above.

DOF Of The Data

The DOF analysis tries to determine the number of truly independent data under certain system geometry. A conventional way to find the DOF is to find the correlation matrix $[\Gamma]$ of the point spread function of the system and diagonalize it. Roughly speaking, the rank of $[\Gamma]$ represents the DOF.

In order to easily apply the results in [10] directly to our radar data, let's normalize the spatial extent of the target as in [10]. If the physical maximum and minimum radio frequency are f_{\min} and f_{\max} corresponding to the 2 way wavelengths

$$\lambda_{\max} = \frac{1}{f_{\min}} \times \frac{C}{2} = \frac{1.5 \times 10^8 \text{ m/sec}}{f_{\min}}$$

$$\lambda_{\min} = \frac{1}{f_{\max}} \times 1.5 \times 10^8 \text{ m/sec} \quad (11)$$

then the normalized frequencies will occupy the spectral ring from

$$f_1 = \frac{a}{\lambda_{\max}} \quad \text{to} \quad f_2 = \frac{a}{\lambda_{\min}}$$

as in Fig. 5.

From [10] the (i, m, k, l) th entry $r_{k,l}^{i,m}$ of $[\Gamma]$ is

$$r_{k,l}^{i,m} = \int_R \int \exp[-j2\pi u_k (\xi \cos \theta_i + \eta \sin \theta_i)]$$

$$\exp[+j2\pi u_l (\xi \cos \theta_m + \eta \sin \theta_m)] d\xi d\eta$$

$$= \frac{J_1[2\pi(u_k^2 - 2u_k u_l \cos(\theta_i - \theta_m) + u_l^2)]^{\frac{1}{2}}}{[u_k^2 - 2u_k u_l \cos(\theta_i - \theta_m) + u_l^2]^{\frac{1}{2}}} \quad (12)$$

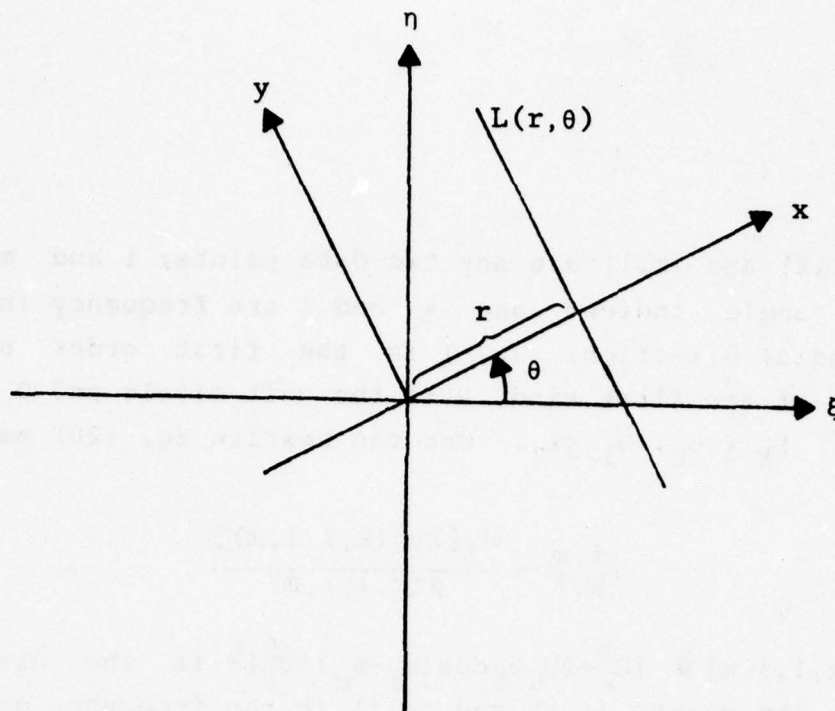


Fig. 4. Projection imaging geometry

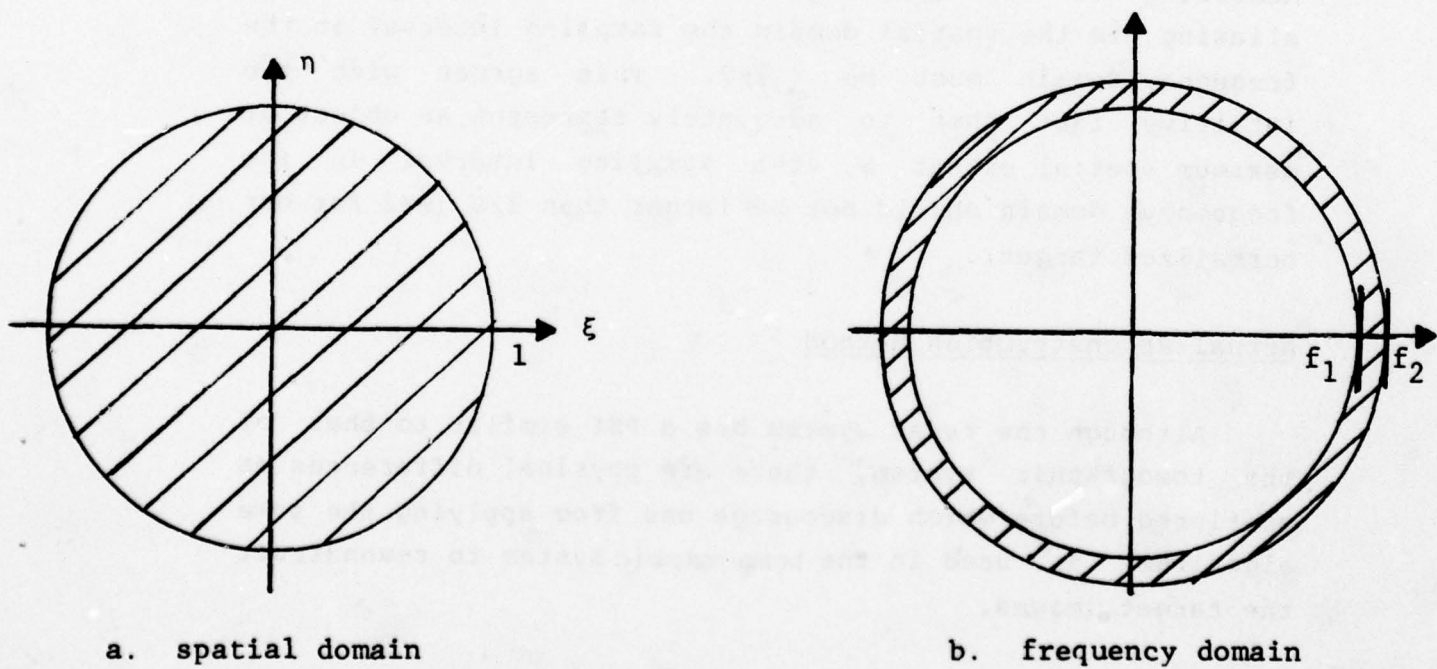


Fig. 5. spatial domain and frequency domain of a spatially normalized target

where (i,k) and (m,l) are any two data points; i and m are azimuth angle indices and k and l are frequency indices along radial direction. $J_1(\cdot)$ is the first order Bessel function of the first kind, R is the unit circle and $0 \leq \theta_i, \theta_m < 2\pi$. $f_1 \leq U_k, U_l \leq f_2$. One can rewrite Eq. (20) as

$$r_{k,l}^{i,m} = \frac{J_1[2\pi\rho(k,l,i,m)]}{\rho(k,l,i,m)} \quad (13)$$

where $\rho(k,l,i,m) \triangleq [U_k^2 - 2U_k U_l \cos(\theta_i - \theta_m) + U_l^2]^{\frac{1}{2}}$ is the distance between the points (i,k) and (m,l) in the frequency domain. Thus one can think of the data as samples of a stationary field with correlation function $\frac{J_1(2\pi\rho)}{\rho}$, which has a power spectrum (Fourier transform of the correlation function) $\text{circ}(r)$ with radial cutoff frequency 1, as shown in Fig. 6. According to the sampling theorem, in order to avoid aliasing in the spatial domain the sampling interval in the frequency domain must be $\leq 1/2$. This agrees with the intuitive fact that to adequately represent an object of maximum spatial extent S , the sampling interval in its frequency domain should not be larger than $1/S$ ($S=2$ for our normalized target).

Actual Reconstruction Method

Although the radar system has a PSf similar to that of the tomographic system, there are physical differences as mentioned before which discourage one from applying the same algorithms as used in the tomographic system to reconstruct the target images.

In Fig. 7 let (x_1, y_1) and (x_2, y_2) be the coordinates of the same targets at two aspect angles whose difference is ϕ . Similar to Eq.(3) we have

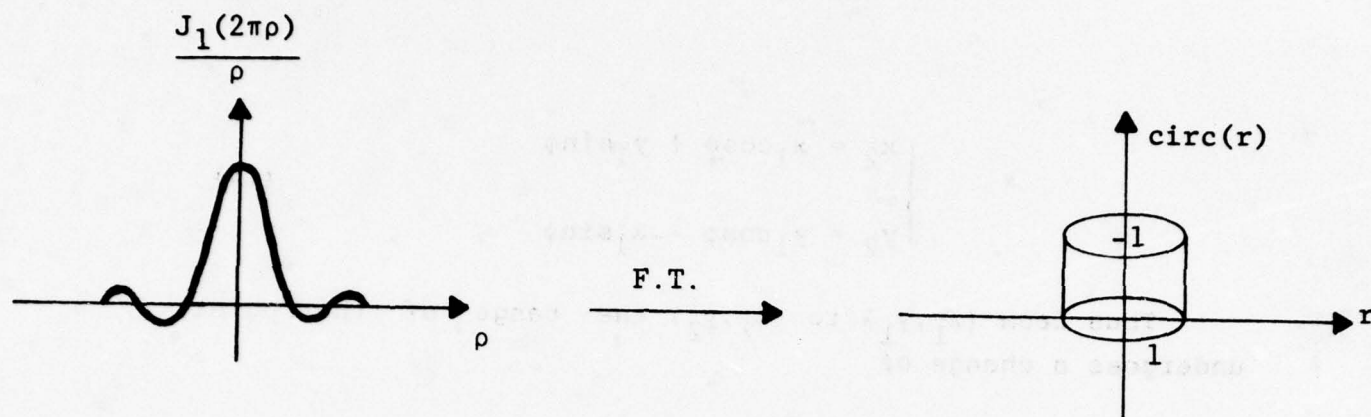


Fig. 6. Correlation of the data field and its Fourier transform

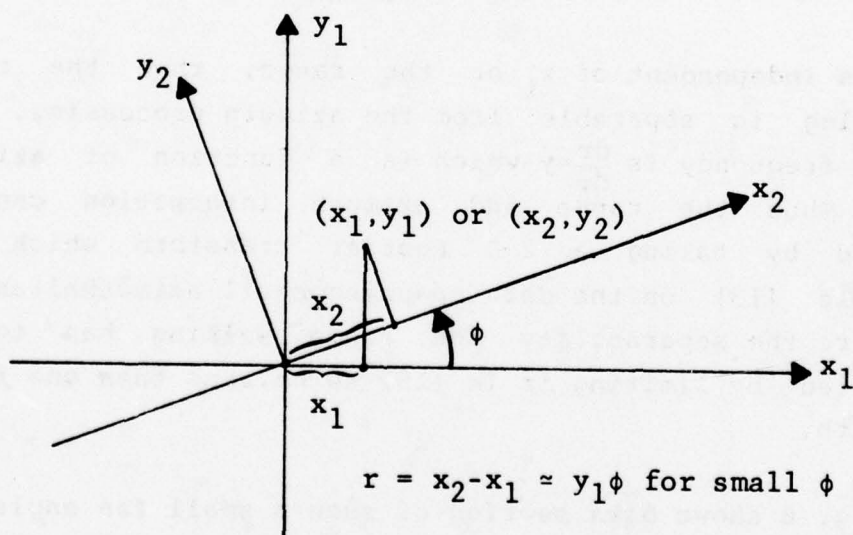


Fig. 7. Separability of azimuthal and range processing at small angle variation ϕ . (Target coordinates fixed on paper)

$$\begin{cases} x_2 = x_1 \cos \phi + y_1 \sin \phi \\ y_2 = y_1 \cos \phi - x_1 \sin \phi \end{cases} \quad (14)$$

Thus from (x_1, y_1) to (x_2, y_2) the range of the point undergoes a change of

$$\begin{aligned} \Delta r(x_1, y_1, x_2, y_2) &= x_2 - x_1 \\ &= x_1 \cos \phi + y_1 \sin \phi - x_1 \\ &= x_1 (\cos \phi - 1) + y_1 \sin \phi \end{aligned} \quad (15)$$

which creates different phase histories to different points. If ϕ is so small that

$$\cos \phi \approx 1 \quad \text{and} \quad \Delta r(x_1, y_1, x_2, y_2) \approx y_1 \sin \phi \approx y_1 \phi \quad (16)$$

which is independent of x_1 or the range, then the range processing is separable from the azimuth processing. The doppler frequency is $\frac{dr}{d\phi} = y$ which is a function of azimuth only. Thus the range and azimuth information can be resolved by taking a 2-D Fourier transform which is separable [13] on the data spanning small azimuthal angle. To insure the separability the range walking has to be prohibited by limiting Δr in (15) to be less than one range bin width.

Fig. 8 shows data section of such a small fan angle ϕ . Let B_r be the range bandwidth with step frequency Δf_r and B_{z1} and B_{z2} be the minimum and maximum azimuth bandwidths at f_{z1} and f_{z2} respectively with step frequencies Δf_{z1} and Δf_{z2} , then

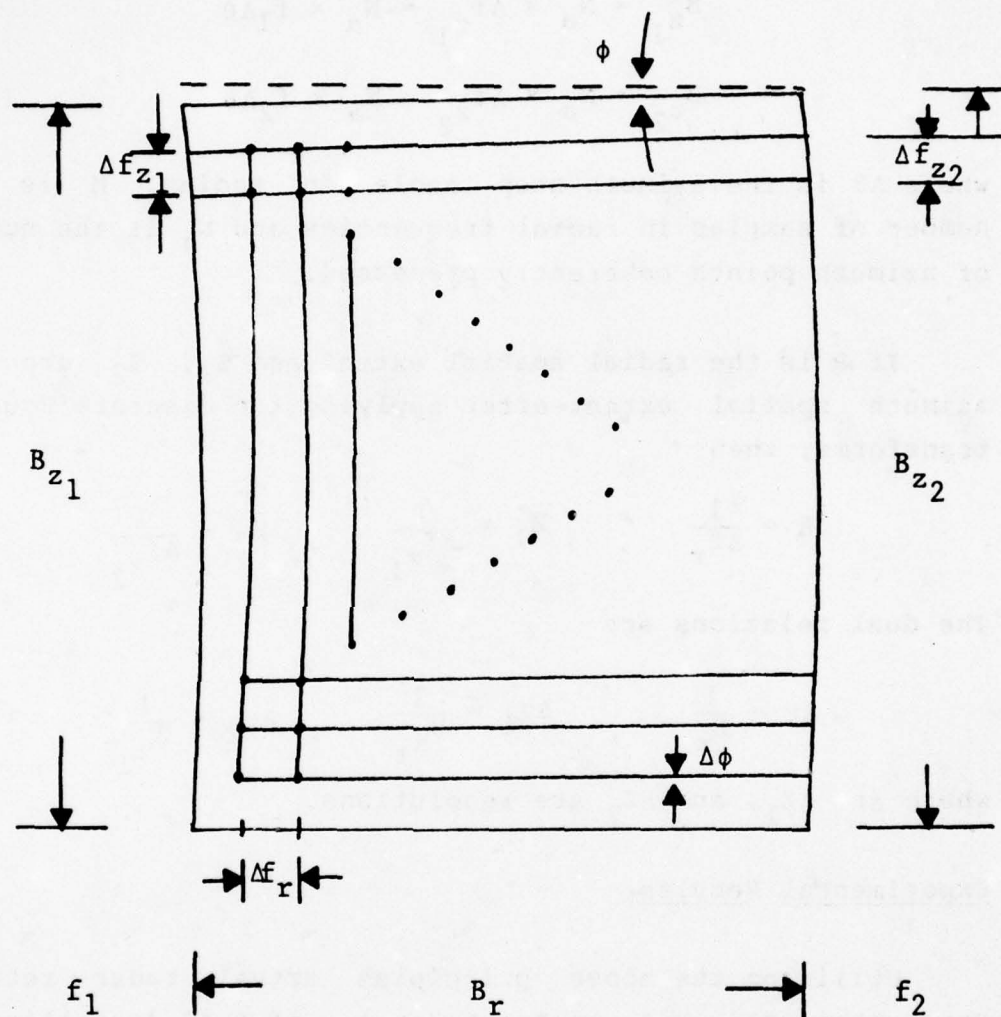


Fig. 8. separable processing of range and azimuth, small fan angle ϕ

$$B_r = M \times \Delta f_r = f_2 - f_1$$

$$B_{z_1} = N_a \times \Delta f_{z_1} = N_a \times f_1 \Delta \theta$$

$$B_{z_2} = N_a \times \Delta f_{z_2} = N_a \times f_2 \Delta \theta \quad (17)$$

where $\Delta \theta$ is the azimuth step angle in radian, M is the number of samples in radial frequencies and N_a is the number of azimuth points coherently processed.

If R is the radial spatial extent and z_1, z_2 are the azimuth spatial extent after applying the discrete Fourier transforms, then

$$R = \frac{1}{\Delta f_r}, \quad z_1 = \frac{1}{\Delta f_{z_1}}, \quad z_2 = \frac{1}{\Delta f_{z_2}} \quad (18)$$

The dual relations are

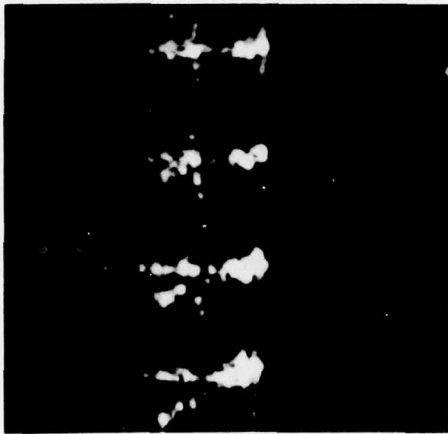
$$\Delta r = \frac{1}{B_r}, \quad \Delta z_1 = \frac{1}{B_{z_1}}, \quad \Delta z_2 = \frac{1}{B_{z_2}} \quad (19)$$

where $\Delta r, \Delta z_1$, and Δz_2 are resolutions.

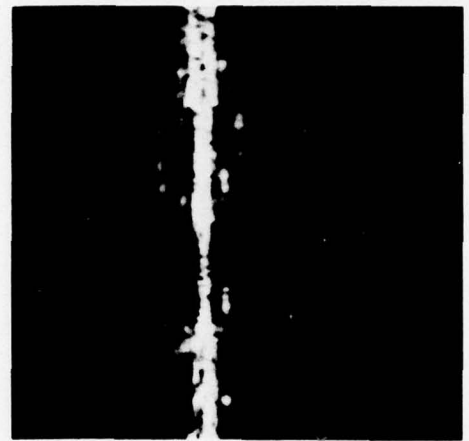
Experimental Results

Utilizing the above principles actual radar returns were processed. A coherent angle of 6.4° (equalling 32 pulses in azimuth) was assumed and 2-D FFT is taken over these pulses. The results are presented in Fig. 9 for various angle of rotation.

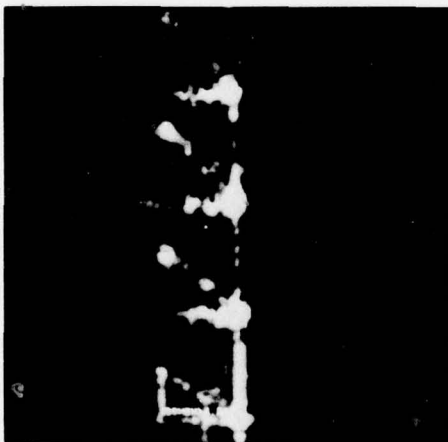
To improve the image quality noncoherent integration is performed with the range cross-range images as in figure 9. With only 7 looks noncoherently summed (at 30° angle intervals) the image of figure 10a results. This is a considerable improvement and clearly shows the outline of



a) 4 looks 0° - 25.6°
(nose)



b) 4 looks 76.8° - 102.4°
(broadside)



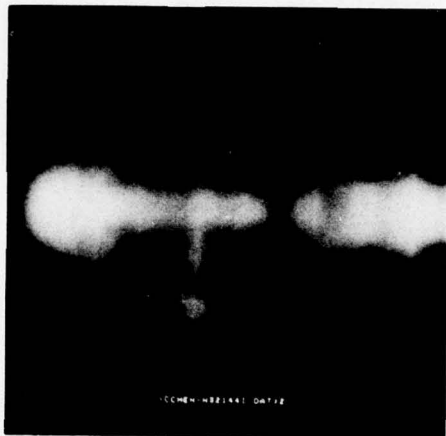
c) 4 looks 128° - 153.6°



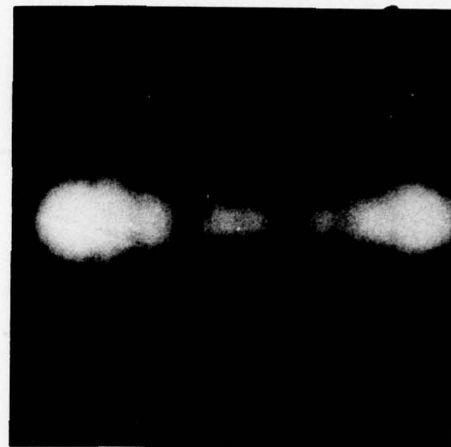
d) 4 looks 153.6° - 179.2°
(tail)

(The Radar is positioned on the right)

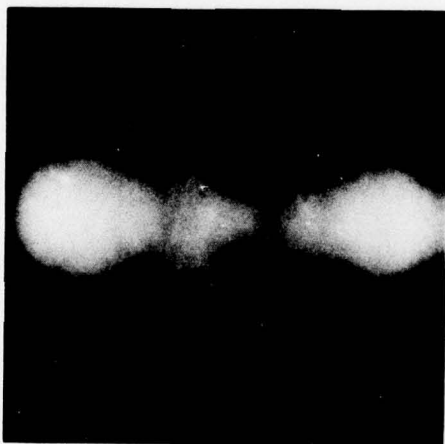
Figure 9. 6.4° coherence in Azimuth at various positions of rotation.



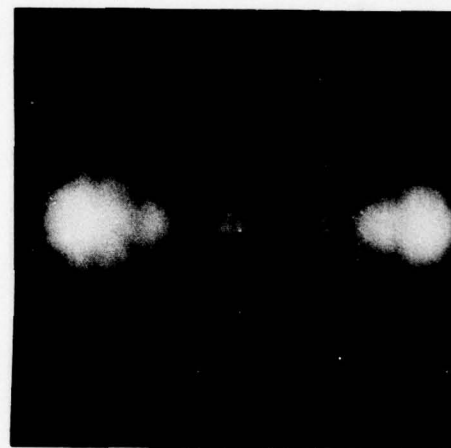
a) 7 looks (6.4° coherence)
spaced 30° apart



b) 28 looks (6.4° coherence)
spaced 6.4° apart



c) 56 looks (3.2° coherence)
spaced 3.2° apart



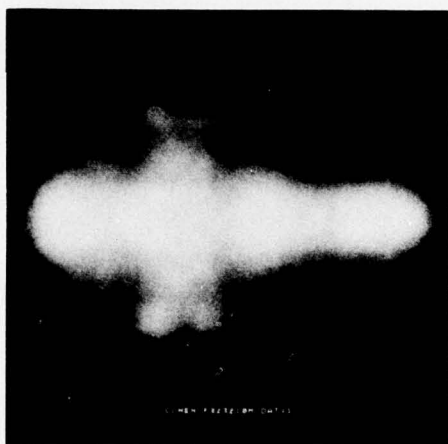
d) 14 looks (12.8° coherence)
spaced 12.8° apart

Figure 10. F102 Imaged for various parameters

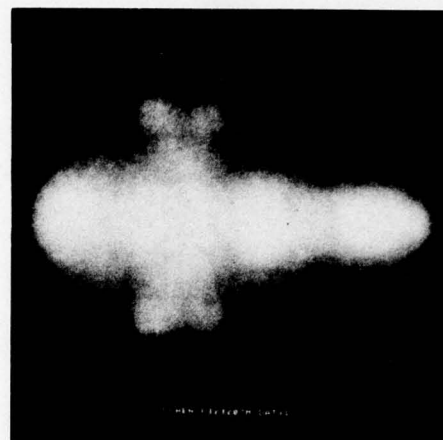
the characteristic delta wing of the F102 aircraft. By noncoherently integrating 28 looks one obtains the results of figure 10b in which a more clear image results. To investigate the degree of coherence necessary (and allowable before "range walking" occurs) figures (10c) and (10d) present result for 3.2° coherence angles and 12.8° coherence angles. In both cases the aircraft is still clearly visible although a certain amount of degradation is beginning to be apparent in both cases.

A second aircraft was imaged using the same parameters as developed above. This aircraft was an F5E and is shorter with stubby wings and wingtip pontoons. Figure 11 presents the results of imaging this airframe with 28 looks at a cross range coherence integration of 6.4° . In this figure the sensitivity to knowledge of the center of motion (rotation in our case) is investigated. This amounts to adjusting the linear phase factors described earlier. By visual inspection probably the best "focused" image is that of figure 11b. The advantage of using the techniques outlined in this figure is that iterative focus can be developed for visual inspection on the stored one dimensional radar returns.

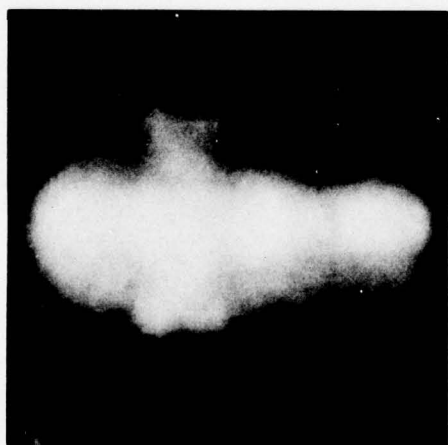
The final figure (figure 12) presents a summary of photographs for the F5E and F102 airframes for both azimuth and elevation plots. Because all parameters are fixed for these images, scales are preserved. Consequently it is clear that the F5E is a smaller aircraft and naturally has a different azimuth and elevation projection than does the F102.



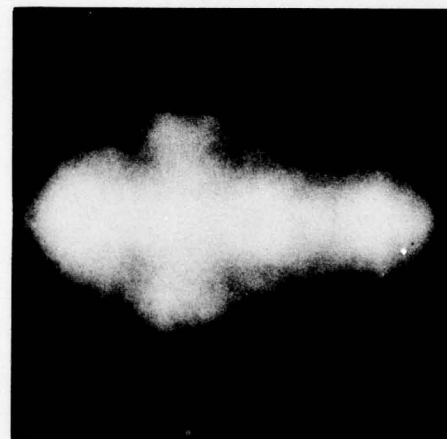
a) 28 looks (6.4° coherence)
centered at range bin 80



b) 28 looks (6.4° coherence)
centered at range bin 79.75

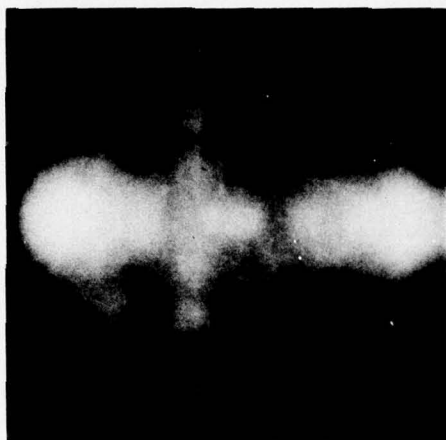


c) 28 looks (6.4° coherence)
centered at range bin 79.5



d) 28 looks (6.4° coherence)
centered at range bin 79

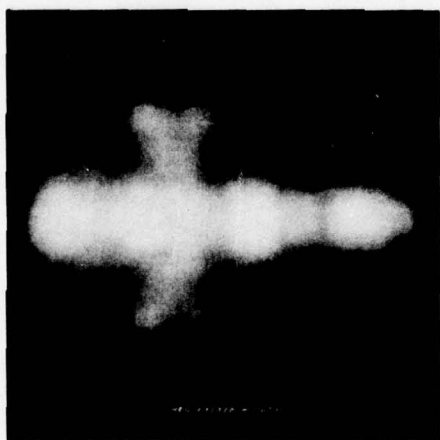
Figure 11. F5E and Center of Rotation



a) F102 Azimuth Image
(28 looks)



b) F102 Elevation Image
(56 looks) (1/2 scale)



c) F5E Azimuth Image
(28 looks)



d) F5E Elevation Image
(56 looks) (1/2 scale)

Figure 12. F102 and F5E Azimuth and Elevation
Images (6.4° Coherence)

Conclusion

This paper has attempted to present the theory of high range resolution radar imaging from both a radar systems viewpoint and a degrees of freedom or numerical analysis viewpoint. Similarity with the computer aided tomographic scanner imaging technology is pointed out. However the differences between the two systems are emphasized and a radar unique reconstruction algorithm is developed for combined coherent and noncoherent imaging. The actual reconstruction method is explained and experimental results developed to illustrate the theories presented. The pictorial images resulting from the computational procedures are surprisingly recognizable and suggest that these techniques may have some practical application in the future.

References

1. M.I. Skolnik, Radar Handbook, McGraw-Hill, 1970.
2. A.W. Richaczek, Principles of High-Resolution Radar, McGraw-Hill, 1969.
3. R.O. Harger, "Synthetic Aperture Radar Systems", Academic Press, 1970.
4. J.C. Kirk, Jr., "A Discussion of Digital Processing in SAR", IEEE Transactions on Aeronautical and Electrical Systems, Vol. 11, No. 3, May 1975.
5. C.C. Chen, "Synthetic Aperture Radar and Imaging System of the Stripping Mode", University of Southern California, Image Processing Institute, USCPI Report 770, Sept. 1977.
6. S. Twomey, "Information Content in Remote Sensing", Applied Optics, Vol. 13, No. 4, 1976.

7. H.C. Marlow, et.al., "The RATSCAT Cross-Section Facility", Proceedings of the IEEE, Aug. 1965.

8. R.N. Bracewell and A.C. Riddle, "Inversion of Fan Beam Scans in Radio Astronomy", Astrophysics Journal, Vo. 150, 1967.

9. L.A. Shepp and B.F. Logan, "The Fourier Reconstruction of a Head Section", IEEE Transactions on Nuclear Science, Vol. NS-21, June 1974.

10. D.G. McCaughey and H.C. Andrews, "Degree of Freedom for Projection Imaging", IEEE Transaction on Acoustics, Speech, and Signal Processing, Vol. ASSP-25, No. 1, Feb. 1977.

11. G.N. Ramachandran and A.V. Lakshminarayanan, "Three Dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions instead of Fourier Transforms", Proc. Nat. Acad. Sci., Vol. 68, 1971.

12. P.R. Smith, et.al., "Image Reconstruction from Finite Numbers of Projections", J. Phys. A.: Math., Nucl. Gen., Vol. 6, March 1973.

13. H.C. Andrews, and B.R. Hunt, Digital Image Restoration, Prentice-Hall, 1977.

3.4 Perceptual Model Coding (Supported by WPAFB under Contract F-33615-77-C-1016)

Charles F. Hall and Harry C. Andrews

The results of the human visual system perceptual model for color image compression represent coding results which are indeed encouraging. The model takes into account the nonlinearities of the brightness response of the eye as well as the necessary spectral and spatial processes to describe non-temporal visual phenomena in a mathematical formulation. Figure 1 presents the basic formulation utilized in this work. Tristimulus values of an image ($R(x,y)$, $G(x,y)$, $B(x,y)$) are passed through a linear transformation T for luminous-like and chrominance-like signals (d_1 , d_2 , and d_3 respectively). The nonlinear logarithm device provides large dynamic range sensitivity and the summing functions thereafter provide color constancy on the chrominance channels e_2 and e_3 . The gain constants C_1 , C_2 , C_3 develop a color space where just noticeable color differences (jncd's) are unit circles. Finally the spatial filters $h_1(x,y)$, $h_2(x,y)$, and $h_3(x,y)$ remove those frequency aspects of an image which the eye does not see as well as emphasizing those aspects to which the eye is particularly susceptible.

This model has been used with considerable success in compressing color images to as low as 1, 1/2, and 1/4 bit per pixel. The color plates accompanying this section present these results in pictorial form.

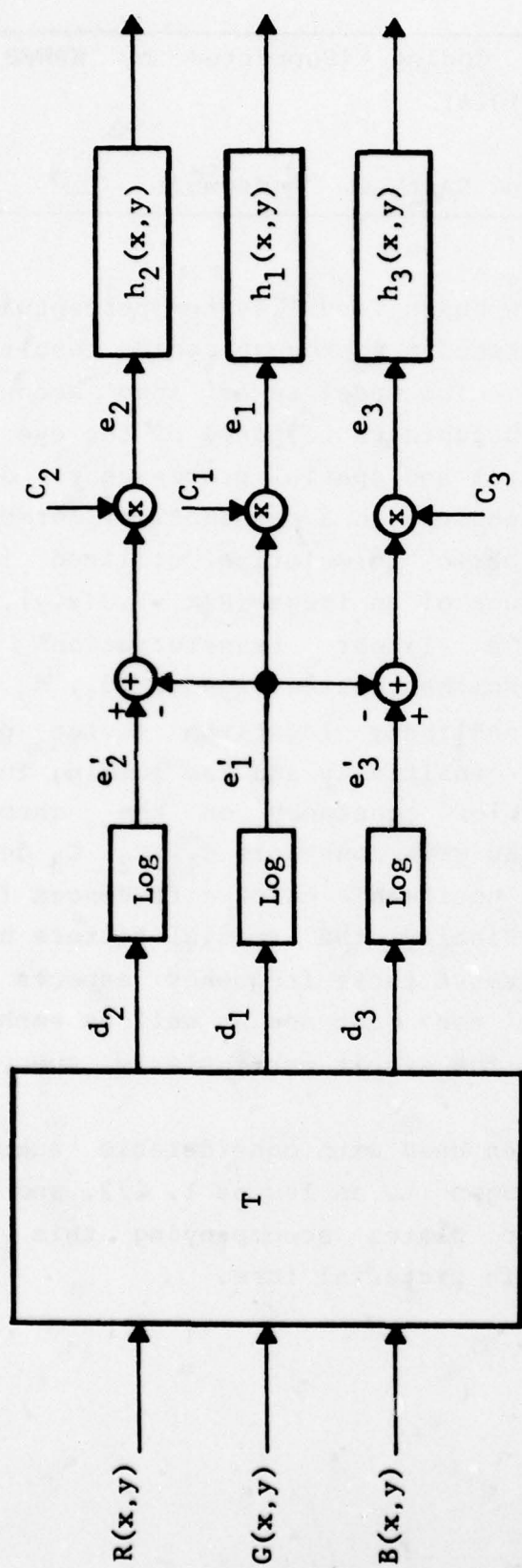


Figure 1. Perceptual Model of Color Vision.



Figure 7.4 Cosine Coded 1 bit/pixel 16 x 16 Blocksize



Figure 7.6 Cosine Coded and Fourier Coded (1 bit/pixel 256 x 256 Blocksize)

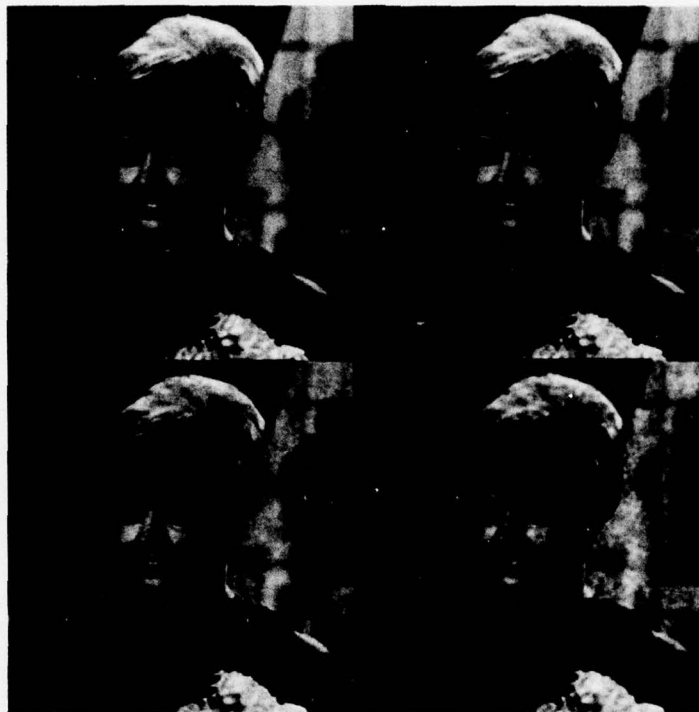


Figure 7.7 Perceptual Power Spectrum Coded ($N = 256$)



Figure 7.8 Perceptual Power Spectrum Coded ANN Image ($N = 512$)



Figure 7.9 Perceptual Power Spectrum Coded LAKE Image (N = 512)

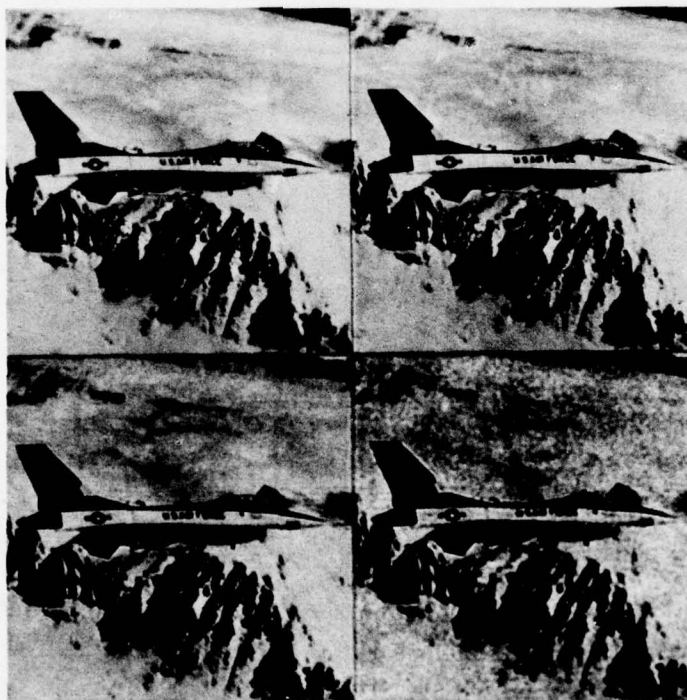


Figure 7.10 Perceptual Power Spectrum Coded F16 Image (N = 512)

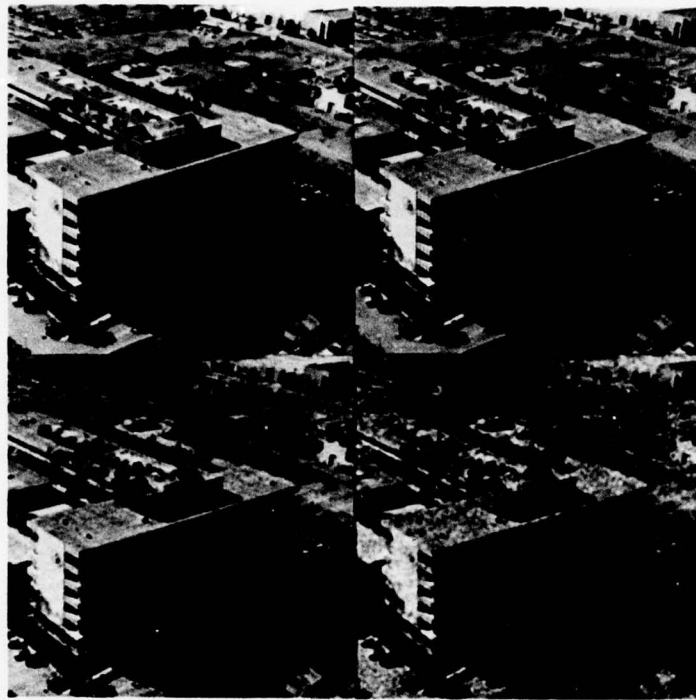


Figure 7.11 Perceptual Power Spectrum Coded BUILDING Image
(N = 512)

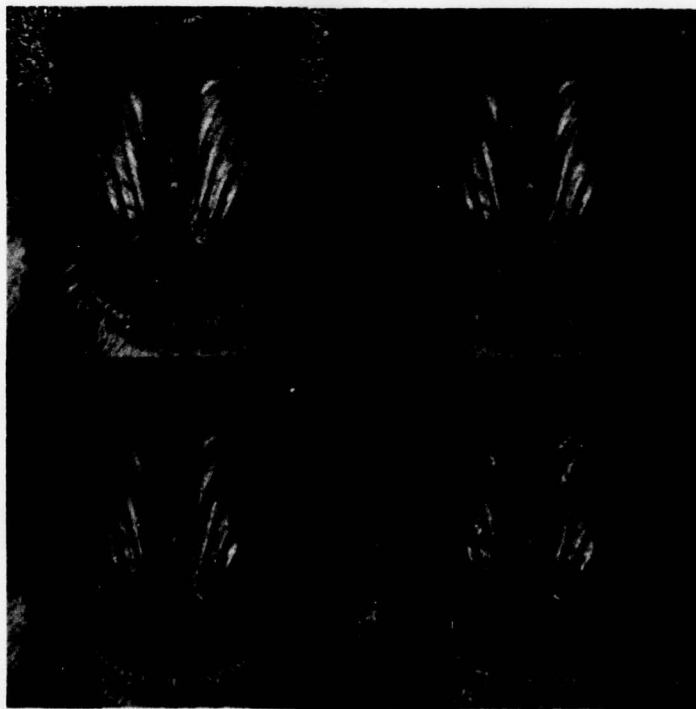


Figure 7.12 Perceptual Power Spectrum Coded BABOON Image
(N = 512)

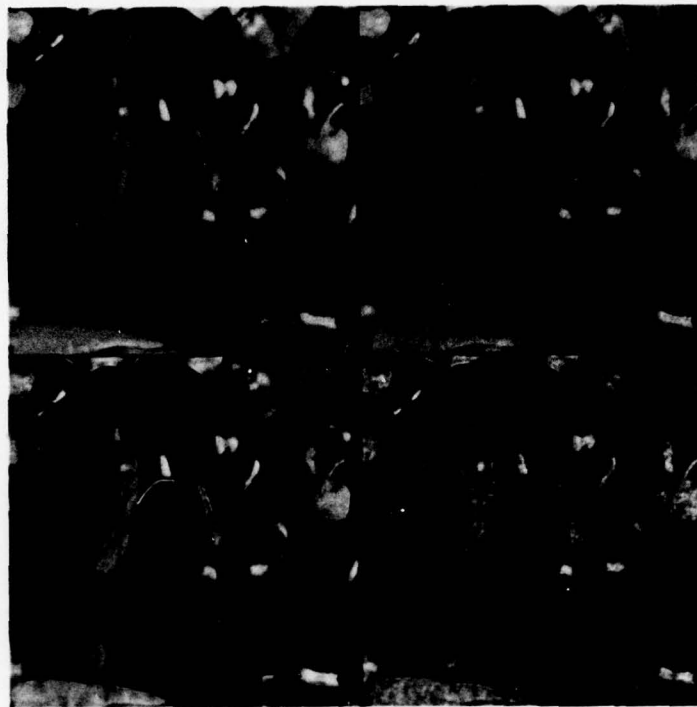


Figure 7.13 Perceptual Power Spectrum Coded PEPPERS Image
(N = 512)



Figure 7.14 Difference Images from ANN Coding Results
(See Figure 7.8)

4. Smart Sensor Projects

This section reports on the successes of the smart sensor CCD chip development under contract at the Hughes Research Laboratories. Earlier reports presented descriptions of two circuits under development, the Sobel circuit and Circuit II which included considerably more complex processes. Both circuits have been demonstrated to produce proper processing results. These past six months have seen the implementation of Circuit II in real time TV (2 MHz reduced horizontal bandwidth) with successful results. It is anticipated that minor redesign will achieve the nominal 10 MHz real time TV bandwidth. In addition progress is being made on other fronts. New circuit designs are in process and a two-year plan is presented in the last contribution to this section.

4.1 Charge Coupled Device Technology for Smart Sensors

Graham R. Nudd, Paul A. Nygaard, and Gary D. Thurmond

Abstract

This paper describes our continuing work [1] to design, fabricate and test charge coupled device (CCD) circuits for image preprocessing. Two test chips containing six processing algorithms have now been fabricated and tested. The processing functions are described together with the circuit implementation and a performance evaluation.

Processor Development

We have completed the design and fabrication of two test chips as shown in Figures 1a and 1b. These circuits

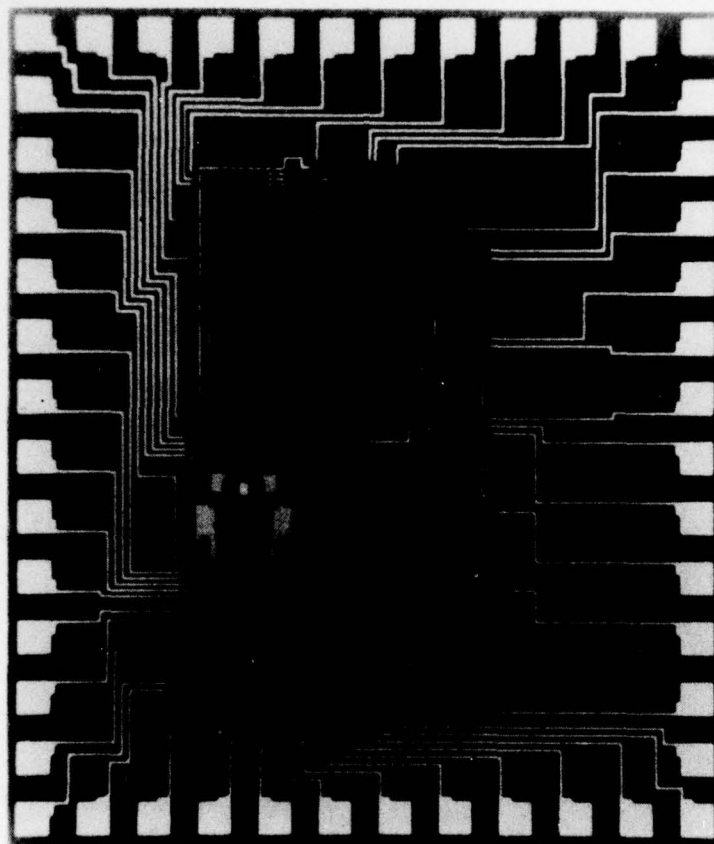


Figure 1a. Photomicrograph of CCD Sobel Circuit.

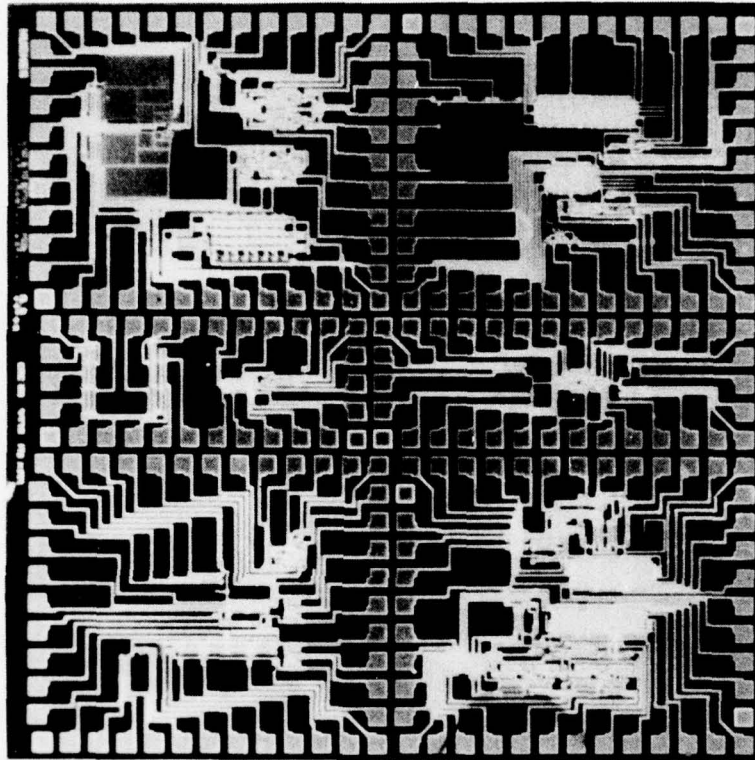


Figure 1b Photomicrograph of Test Circuit II.

are two phase surface channel devices with 8 micron gate lengths. N-type silicon is used to achieve maximum speed. The algorithms implemented are:

Sobel edge detection:

$$f_s = 1/8 \{ |(a+2b+c) - (g+2h+i)| + |(a+2d+g) - (c+2f+i)| \} \quad (1a)$$

$$= |S_x| + |S_y| \quad (1b)$$

Local averaging:

$$f_m = 1/9 (a+b+c+d+e+f+g+h+i) \quad (2)$$

Unsharp masking:

$$f_{usm} = (1-\alpha) e + \alpha f_s \quad (3)$$

Binarization:

$$f_b = \begin{cases} 1 & f_m \leq e \\ 0 & f_m > e \end{cases} \quad (4)$$

Adaptive stretching:

$$f_{as} = \begin{cases} 2 \min |e, r/2| & \text{for } \leq r/2 \\ 2 \max |e, r/2, 0| & \text{for } > r/2 \end{cases} \quad (5)$$

Each is based on a 3 x 3 array of picture elements which are illustrated in Figure 2.

3 × 3 Array

a	b	c
d	e	f
g	h	i

Figure 2. Kernel of Pixels Used in the Calculations, Illustrating the Notations Used in Equations 1 through 5.

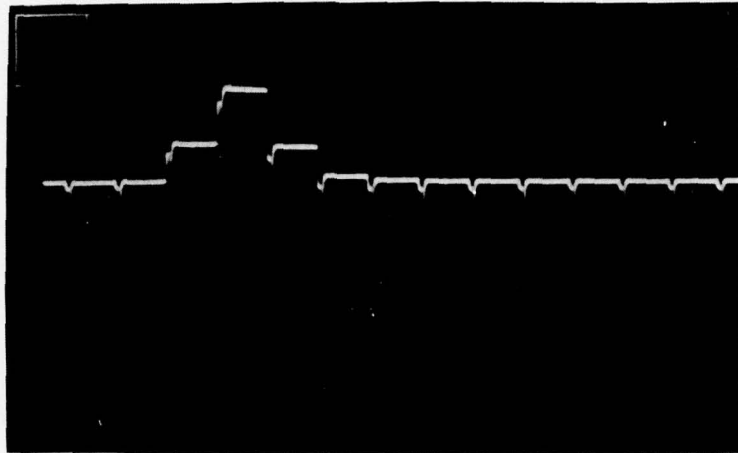
The first circuit, Figure 1a, performs the Sobel operator detecting edges in two dimensions. The processor architecture is arranged in the form of a two-dimensional transversal filter with impulse response for the two edge components.

$$W_x = \begin{bmatrix} 1/8 & 1/4 & 1/8 \\ 0 & 0 & 0 \\ -1/8 & -1/4 & -1/8 \end{bmatrix} \quad (6a)$$

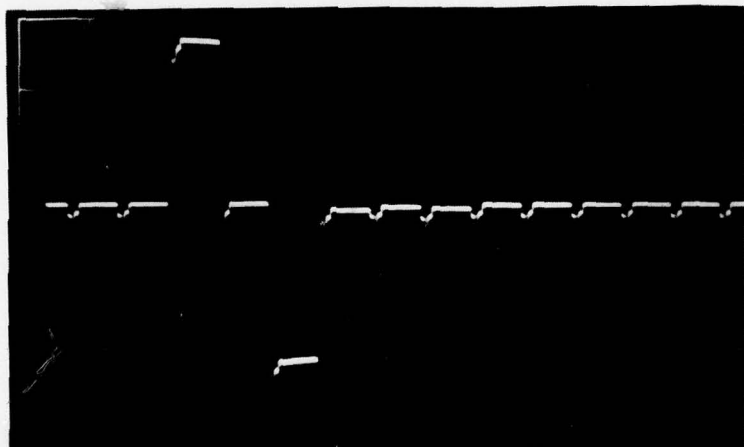
and

$$W_y = \begin{bmatrix} 1/8 & 0 & -1/8 \\ 1/4 & 0 & -1/4 \\ 1/8 & 0 & -1/8 \end{bmatrix} \quad (6b)$$

Using these two components, both the absolute magnitude of the operator, eq.(1), and the edge direction, $\tan \theta = S_x/S_y$ are directly available. The effectiveness of the weighting techniques is shown in Figure 3. The two edge components S_x and S_y then are applied directly to a CCD absolute magnitude

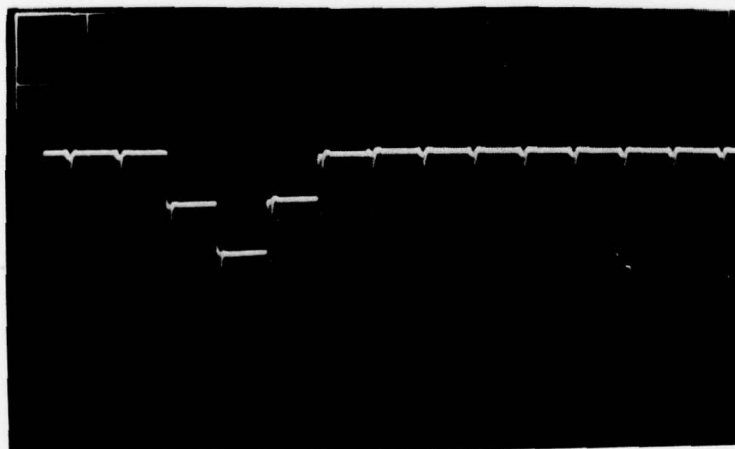


$$S_x = 1/8, 1/4, 1/8$$



$$S_y = 1/4, 0, 1/4$$

(not to same scale)



$$S_y = -1/8, -1/4, -1/8$$

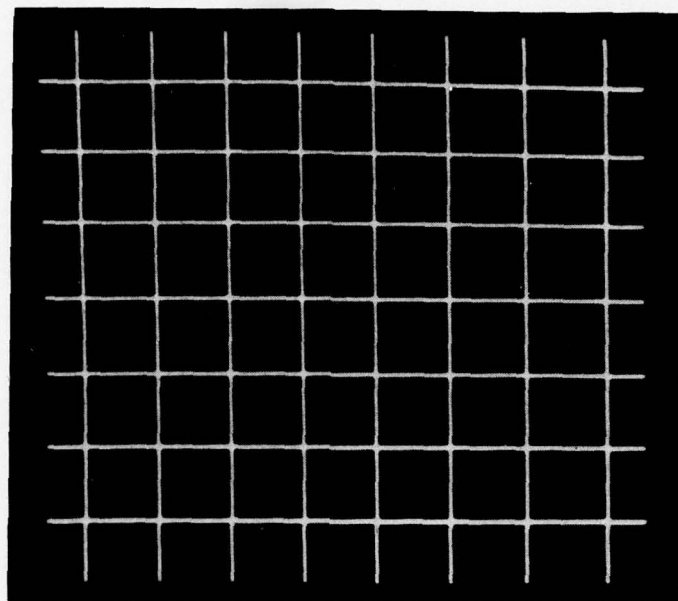
Figure 3. Impulse Response of the 2-D Filter.

operator and a charge summer.

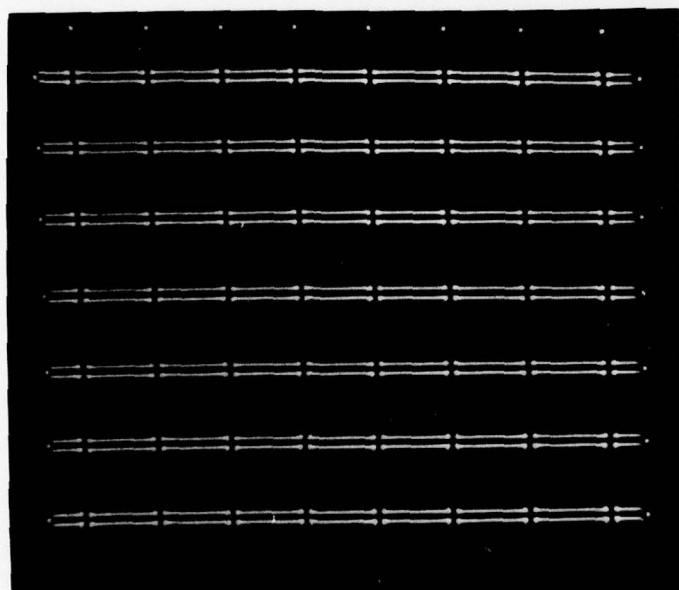
The performance of the CCD edge detector is illustrated in Figure 4 where an original black and white test pattern (4a), the computer simulated Sobel (4b), and the output of the CCD processor (3c) can be compared. The clock rate for this demonstration was 4 kHz, limited primarily by the test facilities. For comparison, the CCD Sobel operation of other optical images is given in Figures 5-7. Our evaluations [2] indicate that at these clock rates, the operation has an accuracy and dynamic range equivalent to four bits. We are currently unable to examine a larger gray scale due to the access time of the processed data from the commercial refresh memory we are using.

We have spent considerable effort developing a real time processing capability to operate the CCD processor from a commercial vidicon camera, the Cohu Model No. 7120 [3]. The basic data rate required for this is approximately 7.5 MHz. We are currently operating our CCD processor at 2 MHz, which results in a slightly unsymmetrical Sobel operation as shown in Figure 8. The frame rate is equivalent to 60 fields/sec with 512 lines as in standard television, however the pixel resolution in the horizontal direction is degraded by approximately a factor of 3. We have tested the circuits at these rates with a variety of images and our intention is to increase the effective data rate in the next phase of the program to achieve truly symmetrical operation.

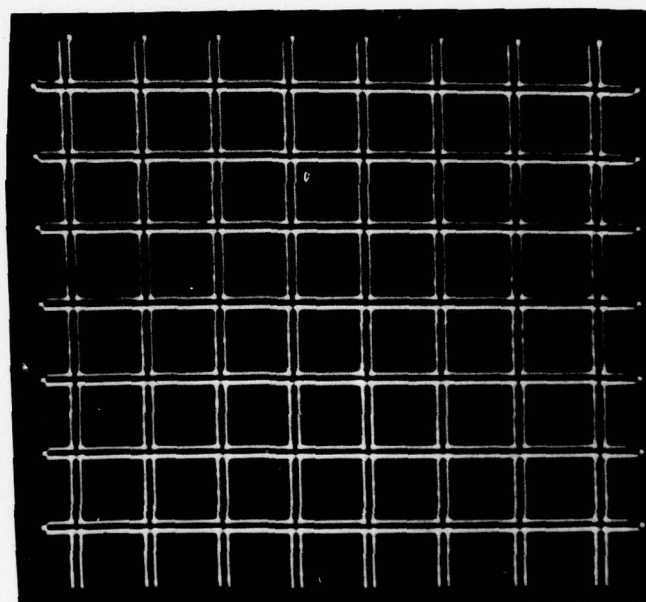
The edge detection circuit described above is basically an important demonstration of two-dimensional nonlinear processing. Our second test chip which performs the operations described in eqs.(1) through (5) is aimed at demonstrating adaptive functions based on the local mean or average. As such, the prime operators are the edge



(a) Original image



(b) Computer simulation

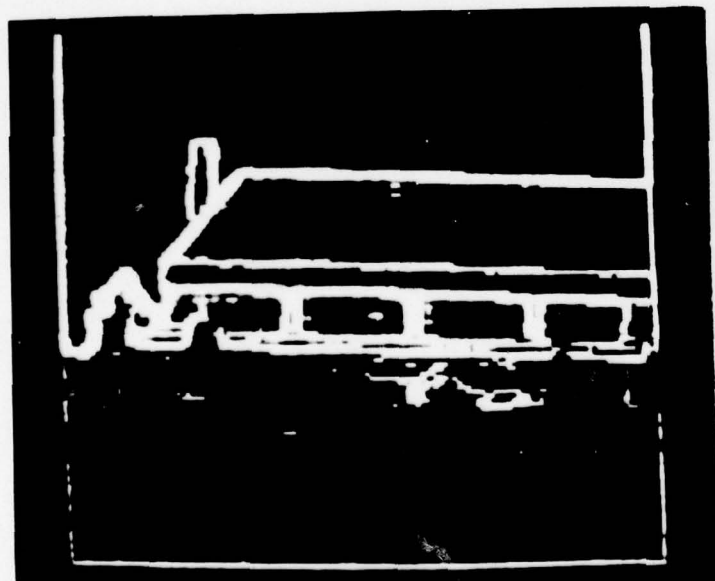


(c) Output for CCD processor

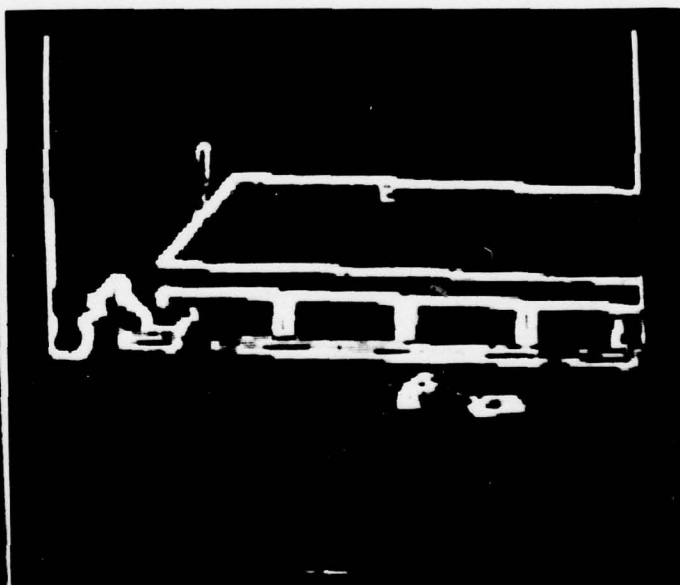
Figure 4. Example of Sobel Edge Detection Using CCD Processor.



(a) Original image



(b) Computer simulation



(c) Output for CCD processor

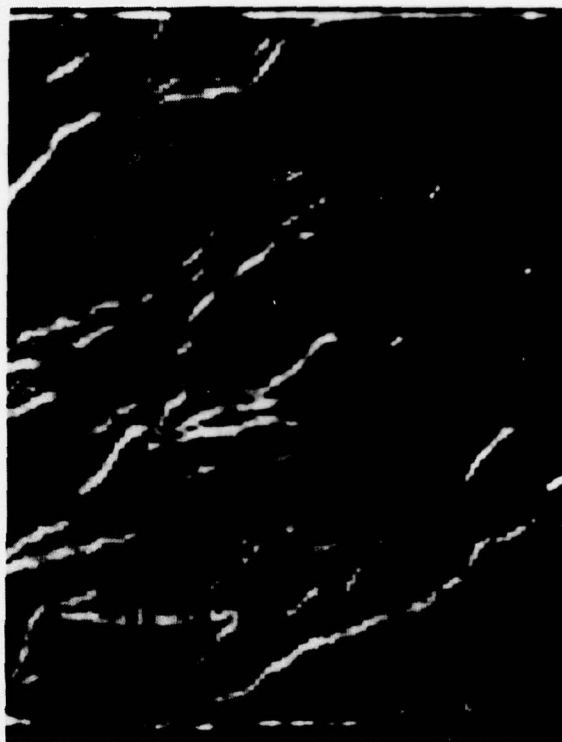
Figure 5. Example of Sobel Edge Detection Using CCD Processor.



(a) Original image

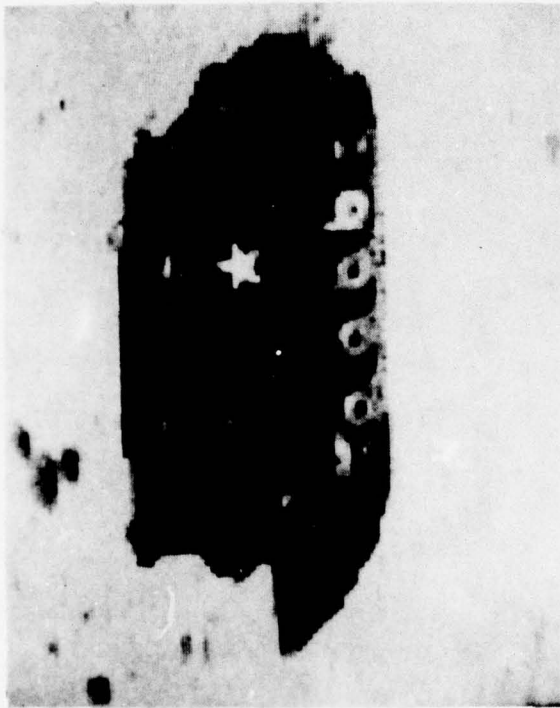


(b) Computer simulation



(c) Output for CCD processor

Figure 6. Example of Sobel Edge Detection Using CCD Processor.



(a) Original image

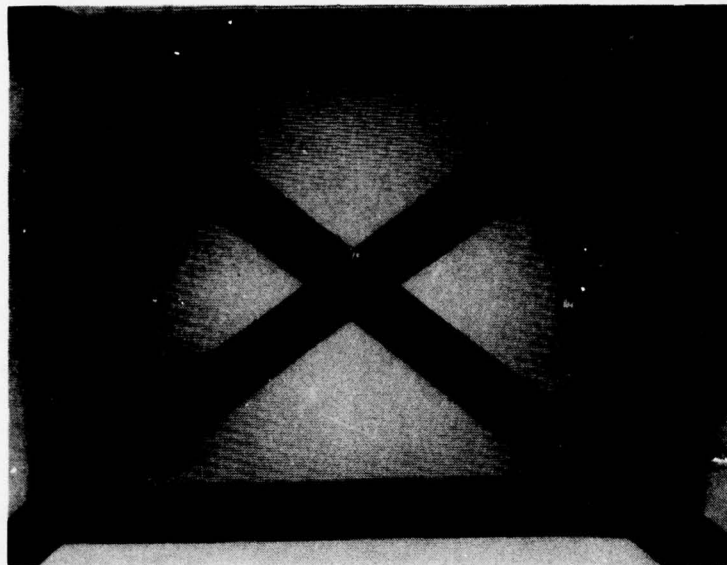


(b) Computer simulation

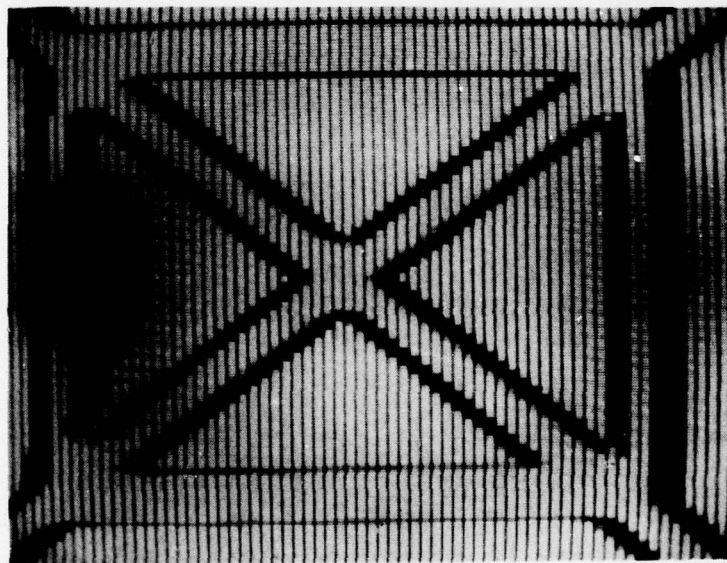


(c) Output for CCD processor

Figure 7. Example of Sobel Edge Detection Using CCD Processor.



(a)



(b)

Figure 8. Example of the Operation of the CCD Sobel Processor Operating in Real-Time from a Commercial Vidicon.

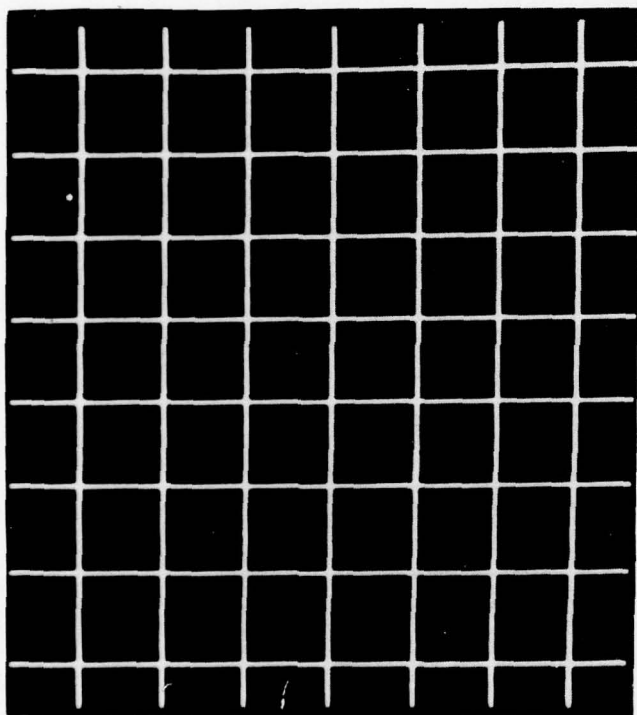
detection, local averaging and the delayed original images. Each of the other algorithms, the unsharp masking, the binarizer and the adaptive stretch are arithmetic combinations of these. The original image, the Sobel and the 3 x 3 mean derived from the second chip are illustrated in Figure 9 for a regular test pattern. Examples of the operation on a true optical image are shown in Figure 9. Each function described in eqs.(1) through (6) (and included in Test Chip II) has been tested and we estimate the overall performance to be equivalent to approximately 4 bits. Testing of linear combinations of the operators described in eqs.(4) through (6) has not been completed at the full video rates, and this effort is currently proceeding. We anticipate no significant problems in this area.

New Concept Development

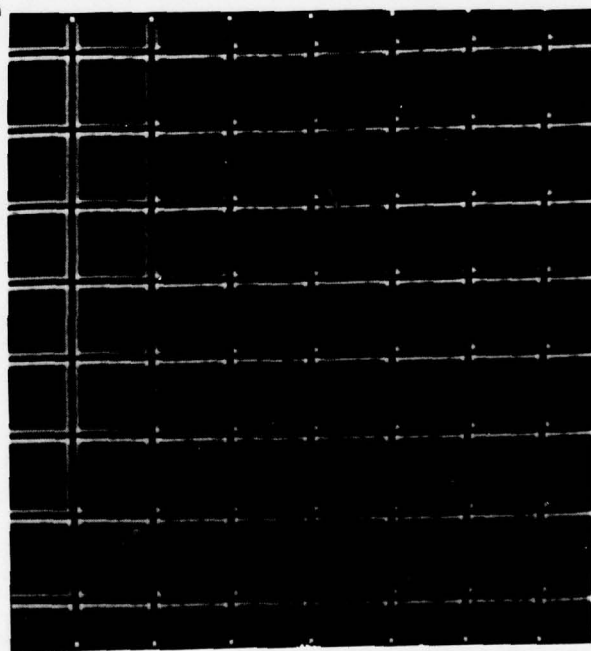
In addition to the above work, we have started concept development and analysis of a third test chip to perform statistics, including a 7 x 7 median filter, an analog histogrammer (including a mode and standard deviation filter, a 7 x 7 programmable processor, and a number of bipolar fixed filters). This work will continue into the next phase of the program when the detailed design, simulation and initial processing will be undertaken.

Development of a Real Time Demonstration Unit

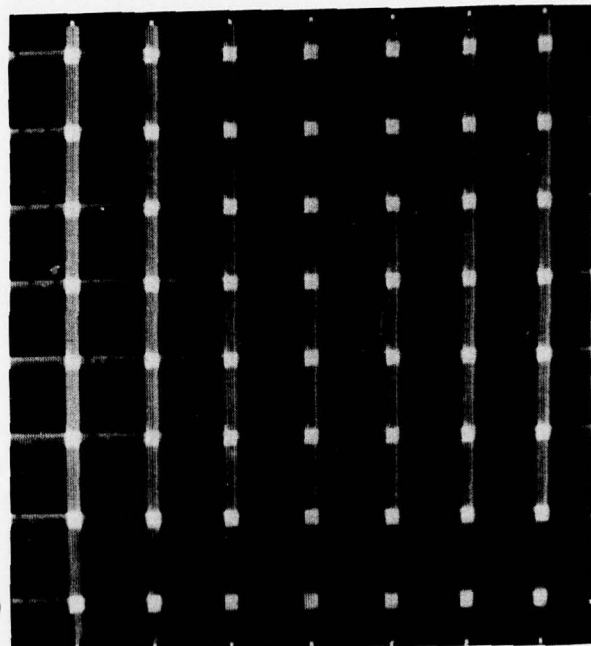
As part of our effort to interface the currently developed processors with a commercial video camera, we are pursuing the development of a small real time demonstration unit which will include the necessary analog CCD delays, the clocks and drivers for our processor, the CCD processors themselves and a small video display unit. This work is well underway, most of the interface circuitry having been



(a) Original image



(b) Sobel edge detection

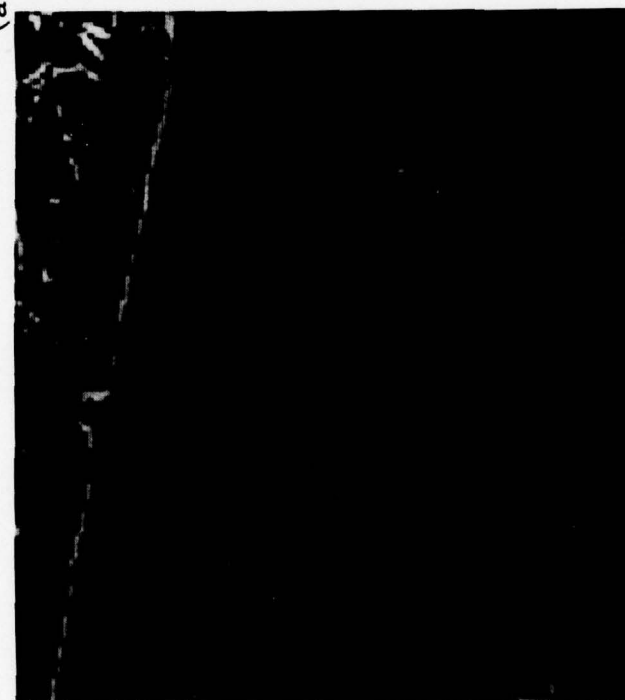


(c) Local 3 x 3 mean

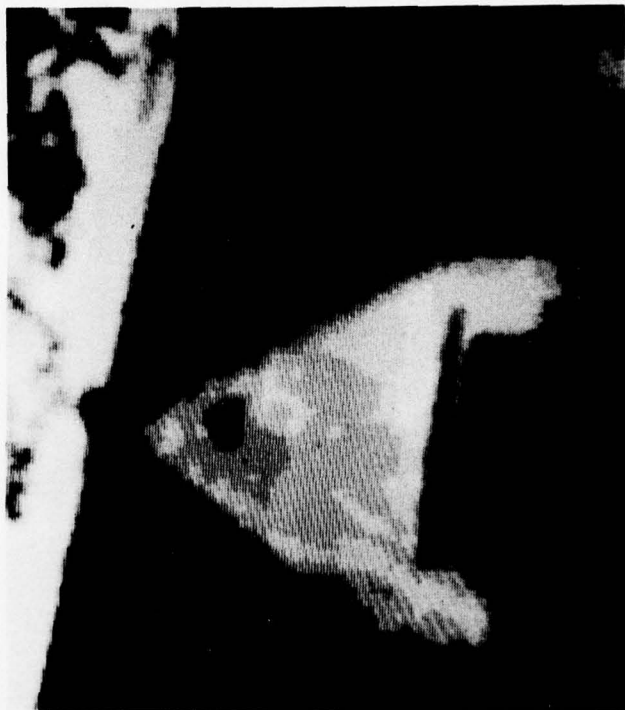
Figure 9. Illustration of Performance of Test Chip II.



(a)



(b)



(c)

Figure 10. Examples of the Operation on a True Optical Image.

designed, and we plan to have the complete unit available in the next phase.

Conclusions

During the previous phase of this program, we have developed CCD integrated circuit processors which perform two-dimensional, nonlinear and adaptive operations at speeds in excess of two orders of magnitude higher than general purpose computers. Our evaluations of this circuit to date indicate that it will perform as predicted [4] and can be interfaced directly to the optical sensors, and as such lead directly to the development of truly smart sensors.

References

1. G.R. Nudd, "CCD Image Processing Circuitry," University of Southern California, USCPI Report 740, March 1977, pp. 142-173.
2. G.R. Nudd and P.A. Nygaard, "Demonstration of a CCD Image Processor for Two-Dimensional Edge Detection," Electronics Letters, Vol. 14, No. 4, February 16, 1978, pp. 83-85.
3. G.R. Nudd, "Chip Helps Detect Targets Automatically," Electronics Magazine, March 16, 1978, pp. 41-42.
4. G.R. Nudd, "CCD Image Processing Circuitry," Proceedings Image Understanding Workshop, Minneapolis, Minnesota, April 1977, pp. 89-94.

4.2 Statement of Work For Follow On CCD Circuitry

Harry C. Andrews

Introduction

This statement of work is designed to outline various research and development tasks to be performed by the Hughes Research Laboratories in cooperation with the USC Image Processing Institute under funding from the Defense Advanced Research Projects Agency (DARPA). It is anticipated that there will be four projects spread over the two years of the contract (October 1, 1977-September 30, 1979) ranging from hardware demonstrations to research studies. The four tasks are delineated below.

Real Time TV Demonstration

At the request of DARPA, task 1 will be to develop a portable self contained television demonstration of the CCD circuits developed on this and past efforts for USCIP and DARPA. Such circuits could include the Sobel, Circuit II, and possibly the 7 x 7 circuits to be concurrently developed over this current two year effort. It is anticipated that the first real time TV demonstration unit be available nominally 6 months after initiation of this contract. However in the interest of simplicity, real time TV can be interpreted to mean processing every other field at regular TV rates thereby implementing the chip operation on nominal 262 x 525 image planes. However it is recommended that no scan rates be changed in order that flicker and other artifacts not be developed.

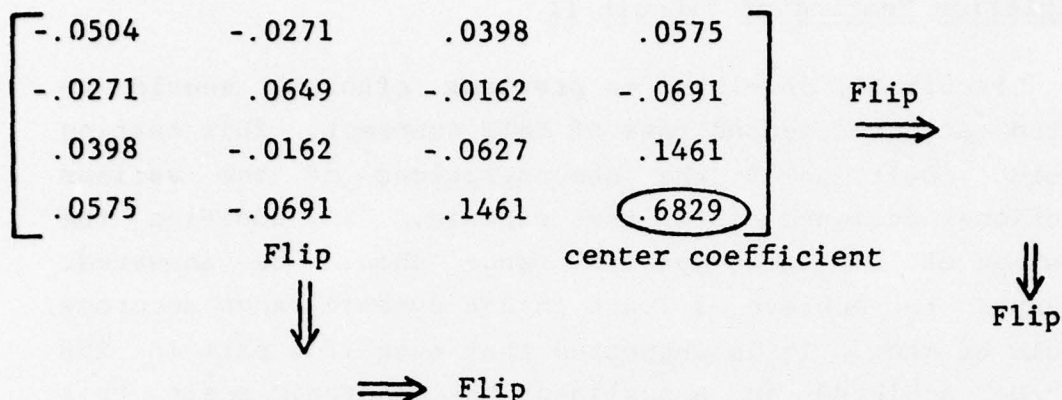
Completion Testing of Circuit II

Circuit II, developed on previous efforts, should be tested as the second task of this contract. This testing should result in 5 kHz demonstrations of the various functions designed into the circuit. In addition the question of SNR and dynamic range should be answered. Attempts to achieve 1 part in 256 dynamic range accuracy should be made. It is suggested that even if 1 part in 256 can be achieved on a nonlinear but monotonic scale, this will be a significant advance. Nonlinear computer generated test patterns are suggested for this experiment.

7 x 7 Circuits

This task will probably require the greatest amount of CCD innovative design capability and probably represents the heart of the follow-on effort. A series of circuits, each 7 x 7 pixels in extent, is to be designed fabricated and tested (certainly at 5 kHz and hopefully at TV rates). A variety of 7 x 7 functions are to be implemented to achieve a) fixed bipolar linear filtering, b) operator programmable bipolar linear filtering, c) data programmable linear filtering, and d) texture related 7 x 7 parameter development.

a) The fixed bipolar linear filtering circuit will be designed to test both the ability to implement bipolar coefficient processes and to measure the accuracy of fixed coefficient numeric values. The filter suggested will be biquadrant symmetric defined by 16 degrees of freedoms as:



b) The operator programmable bipolar linear filter is designed to test the capability of switching filters in the middle of a processing mission. Electronic signals are anticipated for use in this switching process. The above filter could be one of the circuits while another circuit of interest might be another biquadrant symmetric filter defined by:

$$\left[\begin{array}{cccc}
 -.0039\alpha & -.0078\alpha & -.0117\alpha & -.0156\alpha \\
 -.0078\alpha & -.0156\alpha & -.0234\alpha & -.0312\alpha \\
 -.0117\alpha & -.0234\alpha & -.0312\alpha & -.0469\alpha \\
 -.0156\alpha & -.0312\alpha & -.0469\alpha & (1-.0625\alpha)
 \end{array} \right]$$

Let α be a controllable parameter (set to unity for this circuit).

c) The data programmable linear filter is one which changes its structure as a function of the current data within the 7 x 7 window under processing. The observant reader will note that the above filter is unipolar (except for the center pixel). If the unipolar portion is implemented separately, then the size of α can be controlled by the unipolar (low pass) portion of the filter. Letting α be a monotonic function of this low pass signal then will cause greater unsharp masking (edge enhancement) in the brighter regions and less edge enhancement in the less bright regions.

d) The texture related 7 x 7 circuit need not have the precision in coefficient definition given by the above filters. However, what the texture circuit lacks in coefficient accuracy, it more than makes up for in cleverness, variety and nonlinear on chip decisions. In order that as much creativity and unique CCD design be made available for this texture circuit, a variety of suggested operations are listed below for design study prior to fabrication.

- 1) median filter
- 2) mode filter
- 3) dispersion filter
- 4) standard deviation filter
- 5) checkerboard textures

The median filter outputs the median grey level of the 49 pixels within the 7 x 7 window.

The mode filter outputs the mode grey level of the 49 pixels within the 7 x 7 window.

The dispersion filter outputs the height (amplitude or frequency of occurrence) of the mode of the 49 pixels within the 7 x 7 window.

The standard deviation filter outputs the square root of the variance within the 7 x 7 window. (consider sum of absolute differences rather than squares or square roots).

Checkerboard textures output the value of the correlation of the checkerboard with the data within the 7 x 7 window. There are 8 such checkerboards to be considered. See Figure 1.

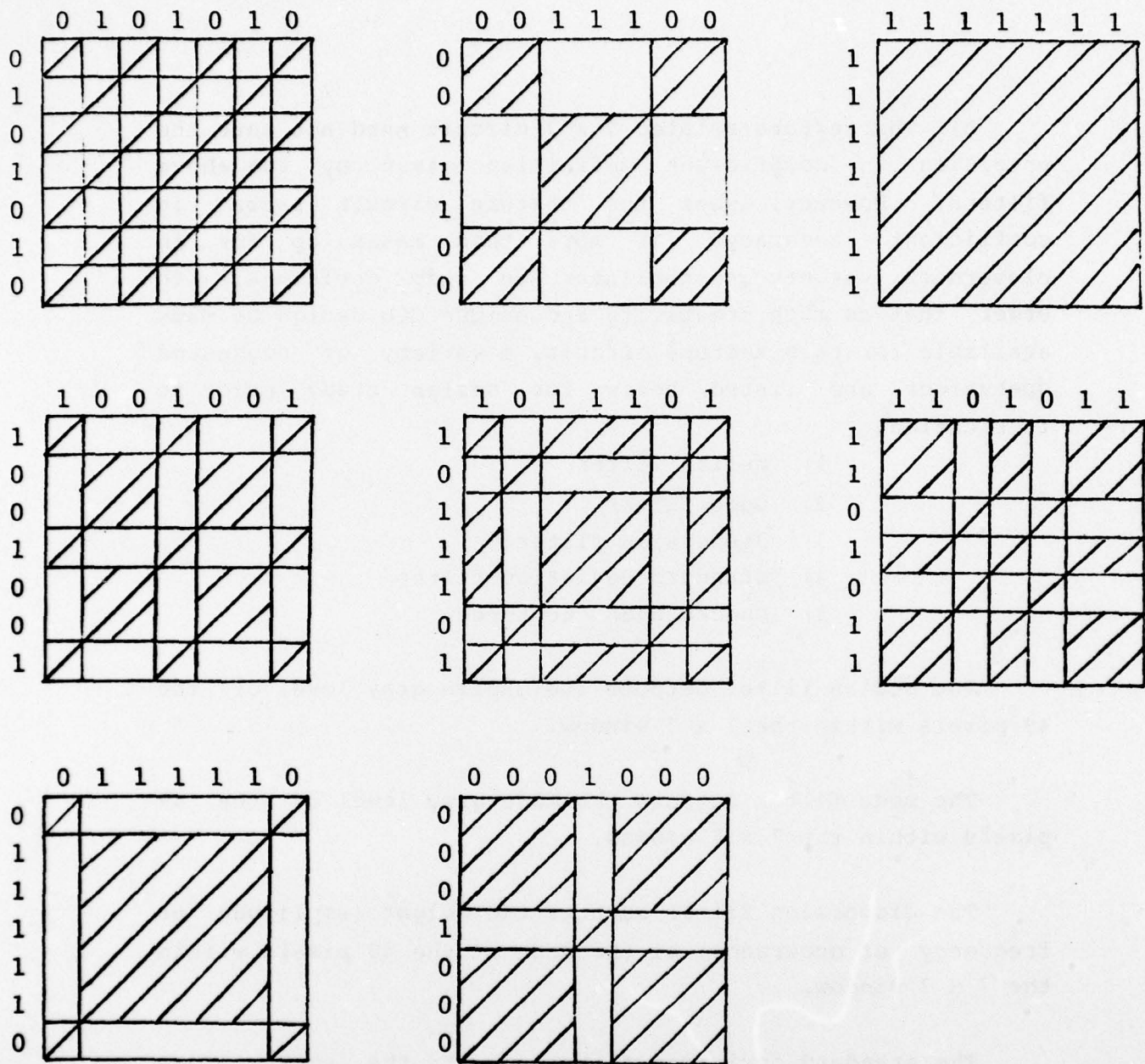


Figure 1. Symmetric separable checkerboard textures

It is hoped that after careful study of these 5 categories of texture circuits, a clever implementation technique will emerge. Certainly a large subset of these suggestions should be implemented in the 7 x 7 texture chip.

Segmentor Study

The final task for this two year effort should be the study of possible implementation of a focal plane (CCD) image segmentor. This should include the combination of outputs of previous circuits and the definition of areas in need of further technological development. This task will be a continually evolving process over the two years and should include dialogue between USCIPI and HRL personnel.

5. Hardware Activities

This section of the semi-annual report is devoted to recent Institute hardware activities dealing with digital imagery equipment.

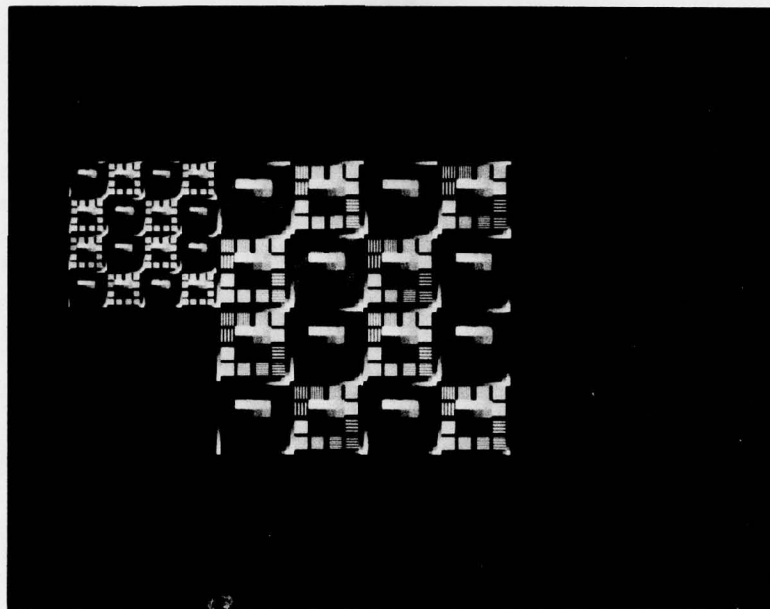
5.1 Hardcopy Acquisition

Harry C. Andrews

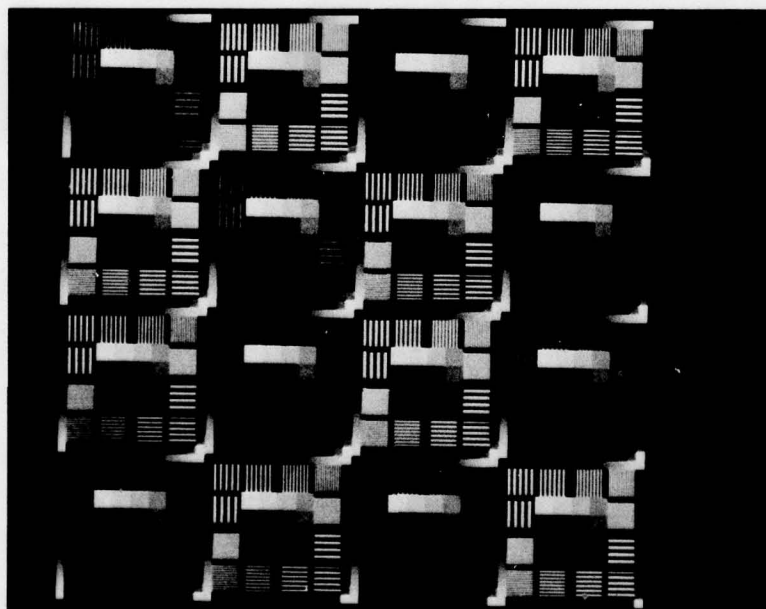
The Image Processing Institute has recently acquired a Dicomed D47 color and monochrome hardcopy device. This equipment now allows us to output high resolution digital imagery (4096 x 4096):

- a) 4" x 5" polaroid
- b) 4" x 5" sheet film
- c) 35 mm. film strips
- d) 35mm. cine film

The device has already proven quite useful and has been calibrated with our own digital test chart. Figure 1 presents this test chart at varying resolution display parameters. Use of the display includes traditional photographic work as well as digitally generated halftone screens for subsequent processing in our hybrid optical-digital laboratory.



(a) 1× and 2× magnification



(b) 4× magnification

Figure 1. Hardcopy Display of USCIP Test Chart.

5.2 The RTTV at ARPA

Harry C. Andrews

The Image Processing Institute has installed a real time television (RTTV) display at ARPA headquarters in Washington, D.C. The device is capable of storing and refreshing a 256 x 256 x 8 bit monochrome image and has an 8 x 8 brightness function memory. The software which supports the system is fairly simple and allows ARPA personnel the ability to access remote image files for up-to-date results of various research projects. Figure 1 presents a picture of the device driving a Sony monitor and our Advent display.

It is expected that CMU will soon have the software for access by ARPA personnel to their image files. Other ARPA contractors wishing to avail themselves of this novel use of the ARPANET are invited to pursue the requirements with their contract monitors and with USCIPi personnel.

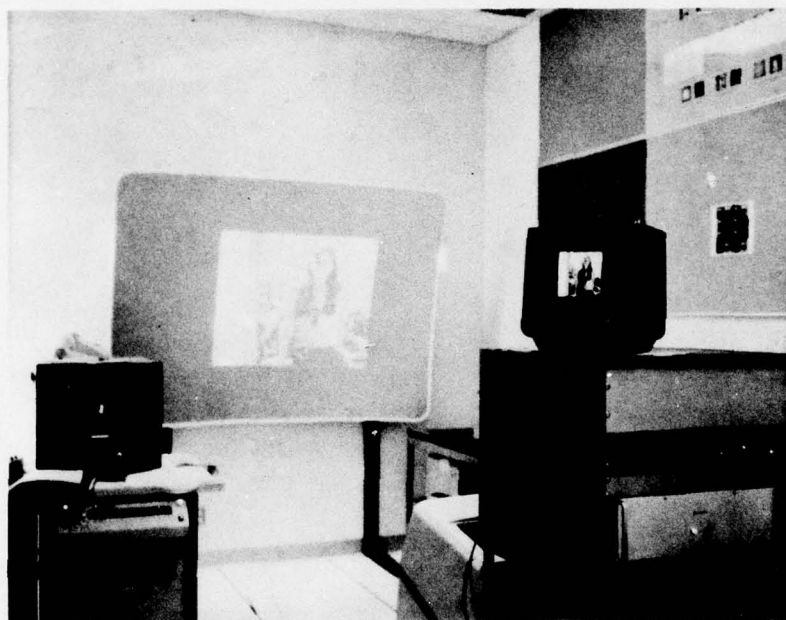


Figure 1. The RTTV Display (before shipment to ARPA in Washington, D.C.)

6. Recent Ph.D. Dissertations

This section includes all dissertations completed since the last reporting period. The one listed here reflects the results of applying the mathematical model of the psychophysics of vision to image compression. Results are extremely encouraging and experimental demonstrations of 24:1 compressions without noticeable degradations are obtainable. The rate distortion analysis also indicates that this model produces results close to the theoretical limit.

Detailed results of this research topic are described in detail in USCIPI Report #790.

6.1 Digital Color Image Compression in a Perceptual Space
(Supported by WPAFB under Contract F-33615-77-C-1016)

Charles F. Hall

Simple mathematical models are developed from the physiological and psychophysical traits of the human visual system. Expressions for the statistical characterization of these models are obtained. When used as a preprocessor, the models are shown to produce images which are statistically compatible with the underlying assumptions necessary to solve the parametric rate-distortion equations. The derived power spectrum equations were used to code black and white and color images with a quality superior to previous results. In addition, it is shown that the preprocessor produces a "perceptual space" in which normalized mean square error is an effective image quality measure.

7. Recent Institute Personnel Publications

1. J. Bescos, I. Glaser and A.A. Sawchuk, "Restoration of Color Images Degraded by Chromatic Aberration," 1977 Annual Meeting, Optical Society of America, Toronto, October 1977, Journal Optical Society of America, Vol. 67, October 1977, p. 1407.
2. J. Bescos and T.C. Strand, "Optical Pseudocolor Encoding of Spatial Frequency Information," to appear in Applied Optics, August 15, 1978.
3. C.C. Chen and H.C. Andrews, "Numeric-Structural Models of Imaging Systems," submitted for publication in IEEE Transactions on Acoustics, Speech, and Signal Processing.
4. C.C. Chen and W. Frei, "Fast Boundary Detection: A Generalization and a New Algorithm," IEEE Transactions on Computers, Vol. C-26, October 1977.
5. G.B. Coleman and H.C. Andrews, "Image Segmentation by Clustering," submitted for publication in IEEE Proceedings.
6. H.S. Hou and H.C. Andrews, "Cubic Splines for Image Interpolation and Digital Filtering," submitted for publication in IEEE Transactions on Acoustics, Speech, and Signal Processing.
7. D.G. McCaughey and H.C. Andrews, "Image Approximation by Variable Knot Bicubic Splines," accepted for publication in IEEE Transactions on Computers.

8. D.G. McCaughey and H.C. Andrews, "The Continuous-Discrete Model: Least Squares Inverses and Singular Function Expansion," submitted for publication in IEEE Transactions on Information Theory.

9. F. Naderi and A.A. Sawchuk, "Estimation of Images Degraded by Film-Grain Noise," to appear in Applied Optics, Vol. 17, 1978.

10. F. Naderi and A.A. Sawchuk, "Detection of Low Contrast Images in Film-Grain Noise," to appear in Applied Optics, Vol. 17, 1978.

11. R. Nevatia and K. Price, "Locating Structures in Aerial Images," Proceedings of ARPA Image Understanding Workshop, Palo Alto, October 1977.

12. R. Nevatia and K. Price, "A Color Edge Detector and Its Use in Scene Segmentation," IEEE Transactions on Systems, Man, and Cybernetics, November 1977, pp. 820-826.

13. R. Nevatia and K. Price, "Image Understanding," Proceedings of Automated Image Processing Workshop, DOD Joint METSTAT Technical Group, Monterey, California, November 1977.

14. R. Nevatia and K. Price, "Evaluation of a Simplified Hueckel Edge-Line Detector," Computer Graphics and Image Processing, December 1977, pp. 582-588.

15. M.J. Peyrovian and A.A. Sawchuk, "Image Restoration by Spline Functions," Applied Optics, Vol. 17, February 1978, pp. 660-666.

16. M.J. Peyrovian and A.A. Sawchuk, "Restoration of Noisy Blurred Images by a Smoothing Spline Filter," Applied Optics, Vol. 16, December 1977, pp. 3147-3153.
17. M.J. Peyrovian and A.A. Sawchuk, "Discrete Representation of Image Degradation Using Monospline Quadrature Formulae," submitted to IEEE Transactions on Acoustics, Speech, and Signal Processing.
18. K. Price, "Experiments in Symbolic Matching of Images," Proceedings of Conference of Algorithms for Image and Scene Analysis, Asilomar, Pacific Grove, California, February 1978.
19. A.A. Sawchuk, "Artificial Stereo," submitted to Applied Optics.
20. A.A. Sawchuk, "Review of Digital Image Processing," by R.L. Gonzales and P. Wintz, Addison-Wesley, Reading, Massachusetts, 1977, to appear in Optical Engineering, 1978.
21. T.C. Strand and A.F. Turner, "Hybrid Optical-Electronic Processing Applied to Chest Radiographs," to appear in Journal of Applied Photographic Engineering.